Structural insights into oxidation of medium-chain fatty acids and flavanone by myxobacterial cytochrome P450 CYP267B1

Jóźwik, Ilona K; Litzenburger, Martin; Khatri, Yogan; Schifrin, Alexander; Girhard, Marco; Urlacher, Vlada; Thunnissen, Andy-Mark W H; Bernhardt, Rita

Published in:
Biochemical Journal

DOI:
10.1042/BCJ20180402

Publication date:
2018

Structural insights into oxidation of medium-chain fatty acids and flavanone by myxobacterial cytochrome P450 CYP267B1

Ilona K. Jóźwik[1], Martin Litzenburger[2], Yogan Khatri[2Φ], Alexander Schifrin[2], Marco Girhard[3], Vlada Urlacher[3], Andy-Mark W. H. Thunnissen[1 Ψ *] and Rita Bernhardt[2*]

[1] Laboratory of Biophysical Chemistry, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 7, 9747 AG, Groningen, The Netherlands

[2] Department of Biochemistry, Campus B2.2, Saarland University, 66123, Saarbrücken, Germany

[3] Institute of Biochemistry, Heinrich Heine University Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany

[Ψ] Current address: Molecular Enzymology Group, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 4, 9747 AG, Groningen, The Netherlands

[Φ] Current address: University of Michigan, Life Sciences Institute, 210 Washtenaw Ave., Ann Arbor, Michigan 48109, United States

***** To whom correspondence should be addressed.

R. Bernhardt, Institute of Biochemistry, Saarland University, Campus B 2.2, 66123 Saarbrücken, Germany. Tel: +49 681 302 4241, Fax: +49 681 302 4739, E-mail: ritabern@mx.uni-saarland.de
or
A.M.W.H. Thunnissen, Molecular Enzymology Group, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 4, 9747 AG, Groningen, The Netherlands. Tel: +31 50 3634380, E-mail: a.m.w.h.thunnissen@rug.nl

Abbreviations: P450- cytochrome P450 monooxygenase, FA- fatty acid, MYR – myristic acid

Keywords: cytochrome P450, CYP267B1, *Sorangium cellulosum*, adrenodoxin, biocatalysis, fatty acids, flavanone, hydroxylation

Abstract

Oxidative biocatalytic reactions performed by cytochrome P450 enzymes (P450s) are of high interest for the chemical and pharmaceutical industries. CYP267B1 is a P450 enzyme from myxobacterium *Sorangium cellulosum* So ce56 displaying a broad substrate scope. In this work, a search for new substrates was performed, combined with product characterization, and a structural analysis of substrate-bound complexes using X-ray crystallography and computational docking. The results demonstrate the ability of CYP267B1 to perform in-chain hydroxylations of medium-chain saturated fatty acids (decanoic acid, dodecanoic acid and tetradecanoic acid) and a regioselective hydroxylation of flavanone. The fatty acids are mono-hydroxylated at different in-chain positions, with decanoic acid displaying the highest regioselectivity towards ω-3 hydroxylation. Flavanone is preferably oxidized to 3-hydroxyflavanone. High-resolution crystal structures of CYP267B1 revealed a very spacious active site pocket, similarly to other P450s able to convert macrocyclic compounds. The pocket becomes more constricted near to the heme, and is closed off from solvent by residues of the FG helices and the B-C loop. The crystal structure of the tetradecanoic acid-bound complex displays the fatty acid bound near to the heme, but in a non-productive conformation. Molecular docking allowed modeling of the productive binding modes for the four investigated fatty acids and flavanone, as well as of two substrates identified in a previous study (diclofenac and ibuprofen), explaining the observed product profiles. The obtained structures of CYP267B1 thus serve as a valuable prediction tool for substrate hydroxylations by this highly versatile enzyme and will encourage future selectivity changes by rational protein engineering.

Introduction

Cytochromes P450 (P450s / CYPs) are heme-thiolate containing monooxygenases present in all domains of life, with >300 thousand sequences identified so far [1]. Physiologically, P450 enzymes are involved in human steroid metabolism, biotransformation of drugs and other xenobiotics, or biosynthesis of secondary metabolites [2,3]. They catalyze a wide spectrum of oxidation reactions using molecular oxygen and incorporate a single oxygen atom into the organic substrate, with the second oxygen atom being reduced to water [4]. To perform the monooxygenase reaction, P450s require electrons to be delivered in order to generate the highly reactive *ferryl-oxo* species (Compound I) in the P450 catalytic cycle. Thus, they need the help of redox partner proteins which oxidize NAD(P)H and transfer the reducing equivalents to the heme-containing P450 [5,6]. In white biotechnology, P450 enzymes are considered attractive biocatalysts for their ability to insert a single oxygen atom into inert C-H bonds, in a selective and controlled manner under mild conditions. In general, bacterial P450 enzymes are more attractive for biocatalytic conversions than their eukaryotic counterparts, due to their easier handling characteristics, like efficient soluble gene expression in *E. coli*, or higher stability. Most importantly, however, the biocatalytic potential of bacterial CYPs lies in their remarkable ability to oxidize substrates in a highly regio- and stereospecific manner [7-10].

To make better use of the P450 catalytic power and versatility to perform synthetically difficult reactions, an expansion of the panel of functionally and structurally well-characterized microbial P450s is desired. Thus far, only a few bacterial P450s have been identified which naturally possess a broad substrate scope like members of CYP154 family (*e.g.* CYP154E1 from *Thermobifida fusca* XY, [11]), CYP109 family [12,13] and the enzyme CYP116B4 from *Labrenzia aggregate* [14]. More often the natural substrate spectrum of bacterial P450s has been altered by protein engineering strategies to allow conversion of

non-physiological substrates. The most prominent example is the P450 BM3 (CYP102A1) from *Bacillus megaterium*, a highly active fatty-acid hydroxylase, for which the vast availability of structural information facilitated successful engineering towards oxidation of many different classes of chemicals like steroids [15], alkanes [16], alkaloids [17], terpenes [18] or polycyclic aromatic hydrocarbons [19]. Notwithstanding the successes of rational protein engineering, the identification and use of wild-type P450s accepting a broad range of substrates is advantageous, as it will circumvent time-consuming procedures like the necessity to generate massive amounts of mutants and tedious screening.

A recently characterized bacterial P450 with an exceptionally broad substrate scope is CYP267B1 from the myxobacterium *Sorangium cellulosum* So ce56 [20], one of the 21 P450 monooxygenases identified in this bacterium [21]. Without the need for any protein engineering, the enzyme was shown to readily accept different classes of substrates, from small carotenoid-derived aroma compounds or sesquiterpenes [22,23] to sterically challenging molecules like epothilone D [24]. CYP267B1 was also shown to oxidize several structurally diverse drugs, such as diclofenac, ibuprofen, oxymetazoline or repaglinide, as well as antidepressant or antipsychotic drugs like amitriptyline, chlorpromazine, imipramine or promethazine [25,26]. Thus, the wild type CYP267B1 shows high catalytic variability, catalyzing not only aliphatic, allylic and aromatic hydroxylations, but also sulfoxidation or epoxidation [22,26]. Despite the wealth of biochemical data available for this versatile P450, its catalytic potential towards different substrate classes might not be fully explored and its structural properties have not been investigated yet.

In this study, the hydroxylation activity of CYP267B1 towards two new classes of substrates, saturated medium-chain fatty acids (decanoic acid, dodecanoic acid and tetradecanoic acid) and a flavonoid compound (flavanone), was demonstrated for the first time, their binding affinities were evaluated and reaction products identified. Structural data

obtained with X-ray crystallography, supplemented by molecular docking results, allowed us to map the active site geometry of CYP267B1 and to predict the binding modes of currently and previously identified substrates. Overall, our results highlight the vast biocatalytic applicability of CYP267B1 and provide crucial structural insights explaining its broad substrate recognition.

Materials and methods

Chemicals and strains

Isopropyl β-D-1-thiogalactopyranoside (IPTG) and 5-aminolevulinic acid (δ-ALA) were purchased from Carbolution Chemicals (Saarbrücken, Germany). Bacterial media were obtained from Becton Dickinson (Heidelberg, Germany). The *Escherichia coli* (*E. coli*) strains C43 (DE3) and BL21 (DE3) were obtained from Novagen (Darmstadt, Germany). Fatty acids and flavanone were obtained from Sigma Aldrich (Schnelldorf, Germany). All other chemicals were purchased from standard sources and of the highest purity available.

Expression and purification of CYP267B1 and its heterologous redox partners

The construct of pET22b_CYP267B1 encoding the *cyp267b1* gene from *Sorangium cellulosum* So ce56 (including a C' terminal His_6-tag) (GenBank: CAN97336.1) [25] was transformed into *E. coli* C43 (DE3) cells and grown overnight at 37°C in NB-I medium containing 100 µg/mL ampicillin (150 rpm). Subsequently, 2 L baffled Erlenmeyer flasks, each filled with 0.4 L of Terrific Broth (TB) medium (with 100 µg/mL ampicillin), were inoculated (1:100) with that overnight culture, and incubated at 37°C at 90 rpm, until OD_{600} reached 0.8-1. Then the temperature was set to 28°C and 1 mM IPTG and 0.5 mM δ-ALA were added to induce the expression and support heme synthesis, respectively. Cells were

harvested after 48 h by centrifugation for 30 min at 4500 *g*. Collected pellets were then stored at -20°C. Purification procedure started with pellet resuspension in 25 mM Tris (pH 7.4) buffer containing 0.5 mM EDTA, 10 mM NaCl and 1 mM phenylmethylsulfonyl fluoride (PMSF). The cells were lysed by sonication for 15 min on ice, followed by a centrifugation step for 45 min at 75000 *g*, at 4°C. The supernatant was loaded onto Ni-NTA column and washed with 25 mM Tris (pH 7.4) buffer containing 0.1 mM EDTA, 0.1 mM dithioerythritol (DTE) and 5 mM histidine. The elution was performed with the same buffer, but containing 150 mM histidine. Red-colored fractions were pooled and dialyzed against 1 L of 25 mM Tris (pH 7.4) buffer containing 0.1 mM EDTA and 10 mM β-mercaptoethanol. Dialysis buffer was exchanged after 2 h and 16 h, respectively. The protein solution was then concentrated and loaded onto Superdex 75 column equilibrated with 25 mM Tris (pH 7.4) buffer containing 2% glycerol. In order to prepare substrate-bound complexes for crystallization, the buffer used during size-exclusion chromatography step was supplemented with 100 µM of the corresponding substrate (added from a 10 mM stock solution in dimethyl sulfoxide (DMSO)). Finally, intensively red colored fractions showing a spin state shift towards 390 nm in the measured UV-visible absorbance spectrum (or only a small shoulder as in case of flavanone) were concentrated and flash frozen with liquid nitrogen. The pure, aliquoted samples of CYP267B1 were stored at -80°C until further use.

The redox partner proteins used in this work, a truncated form of bovine adrenodoxin ($Adx_{4\text{-}108}$) and adrenodoxin reductase (AdR) were expressed and purified as described previously [27,28].

Substrate binding assay and determination of dissociation constants (K_d)

Substrate-induced spin state shifts were estimated using tandem quartz cuvettes, a double-beam spectrophotometer (UV-2101PC, Shimadzu, Japan) and performing the binding assay as described elsewhere [29]. Purified CYP267B1 (5 µM) was dissolved in 20 mM potassium

phosphate buffer (pH 7.4) and titrated with increasing concentrations of the corresponding substrate from either 1.0 or 10 mM stock solution in DMSO. Difference spectra were recorded from 350 to 500 nm. In order to determine the dissociation constant (K_d), the averaged peak-to-trough absorbance differences (ΔA_{max}) were plotted against increasing substrate concentration. Plots of dodecanoic and tetradecanoic acid were fitted with Origin 8.6 software (OriginLab Corporation, Northampton, MA) by tight binding quadratic equation [$\Delta A = (Amax/2[E])\{(K_d+[E]+[S])-\{(K_d+[E]+[S])^2 - 4[E][S]\}^{1/2}\}$], whereby ΔA represents the peak-to-trough absorbance difference at every substrate concentration, Amax is the maximum absorbance difference at saturation, [E] is the enzyme concentration, and [S] is the substrate concentration. The plot of decanoic acid was fitted by hyperbolic regression [$\Delta A = (Amax[S]/K_d+[S])$] using Origin 8.6 software. All titrations were performed in triplicate.

***In vitro* conversions**

A protein ratio of CYP: Adx: AdR of 1: 10: 3 (for fatty acids) or 1: 20: 3 (for flavanone) was used in a reconstituted *in vitro* system consisting of purified CYP267B1 (0.5 µM), $Adx_{4\text{-}108}$ (5 µM/ 10 µM), AdR (1.5 µM), $MgCl_2$ (1 mM), glucose-6-phosphate (5 mM) and glucose-6-phosphate dehydrogenase (1U) in a final volume of 250 µL of potassium phosphate buffer (20 mM, pH 7.4) was used. The substrates (dissolved in DMSO) were added to a final concentration of 200 µM. The reactions were initiated by addition of 500 µM NADPH. After 30 min at 30°C the reactions were quenched. 500 µL of ethyl acetate (EtOAc) were used twice to extract flavanone and its conversion products. The combined organic phases were evaporated in vacuum and prepared for further analyses. The reactions of the fatty acids were stopped by adding 20 µL of 37% HCl, extracted with 1 mL diethyl ether, dried over anhydrous $MgSO_4$ and the organic phase was evaporated. The residues were dissolved in 40 µL *N*,*O*-bis(trimethylsilyl)trifluoroacetamide containing 1% trimethylchlorosilane and

incubated for 30 min at 80°C before analysis. A negative control with none CYP added was performed for each substrate to verify the P450-dependent reaction.

Whole-cell conversions

The whole-cell reaction for the conversion of flavanone was performed as described previously [26]. The *E. coli* BL21 (DE3) cells were transformed with two plasmids, one encoding CYP267B1 (pET22b_CYP267B1) and one encoding the autologous redox partners from *S. cellulosum* So ce56, FdR_B (Ferredoxin-$NADP^+$ reductase) and Fdx8 (Ferredoxin 8) (pCDF_F8B). After 21 hours of protein expression in 1 L scale (5x 200 mL in 2 L baffled flasks), 200 µM flavanone were added and incubated for 48 hours at 30°C, followed by extraction with the same volume of EtOAc. The extraction step was repeated, the organic phases were pooled and evaporated to dryness. The crude extract was stored at 4°C until further purification.

Product purification was performed by column chromatography using silica gel with a mobile phase consisting of hexane and ethyl acetate (95:5). Fractions were collected and analyzed by thing layer chromatography (TLC) using anisaldehyde staining for visualization. Product containing fractions were pooled, evaporated to dryness and dissolved in $CDCl_3$ for NMR analysis.

Product profile analysis by GC-MS and HPLC

Analysis of the product profile and detailed characterization of hydroxylated fatty acid products was performed via gas chromatography-mass spectrometry (GC-MS) as described previously [12]. Briefly, the dried residue was dissolved in 35 µL of N,O-bis(trimethylsilyl)trifluoroacetamide with 1% trimethylchlorsilan and incubated at 80°C for 30 min to produce derivatized products prior to analysis. The analysis of products was

carried out on a GC-MS-QP2010 (Shimadzu, Tokyo, Japan), equipped with a FS-Supreme-5 column (30 m × 0.25 mm × 0.25 μm; Chromatographie Service, Langerwehe, Germany). A 0.5 μL sample was injected for analysis using a split of 30. The purge flow was 3 mL/min (column flow = 0.61 mL/min), with helium as the carrier gas. The column temperature was maintained at 160°C for 1 min, ramped to 260°C at a rate of 10°C/min, then to 300°C at a rate of 40°C/min and held for 3 min. The hydroxylated fatty acid products were identified by their specific fragmentation pattern in the silylated form, in which mainly two major fragments were considered.

The reaction mixtures resulting from flavanone conversions were analyzed via HPLC. The system used combined a PU-2080 HPLC pump, an AS-2059-SF autosampler, a MD-2010 multi wavelength detector (Jasco, Gross-Umstadt, Germany) and a Nucleodur 100–5 C_{18} column (5 μM, 4.0 x 125 mm, Macherey–Nagel, Düren, Germany) thermostated at 40°C. Mobile phases consisted of 10% (v/v) acetonitrile in water (A) and pure acetonitrile (B). For analysis, a linear gradient from 10 to 80% of B (for 30 min) was used, followed by a linear gradient to 100% B for 10 min. Then the column was equilibrated with 10% of B for 5 min. The flow rate was set to 1 mL/min.

NMR analysis

NMR spectra were recorded with a Bruker DRX 500 (Rheinstetten, Germany) NMR spectrometer. A combination of ^{1}H, ^{13}C, ^{1}H, ^{1}H-COSY, HSQC and HMBC was used to elucidate the structure. All chemical shifts are relative to $CHCl_3$ (δ=7.24 for ^{1}H NMR) or $CDCl_3$ (δ=77.00 for ^{13}C NMR) using the standard δ notion in parts per million (ppm).

Crystallization

The samples of CYP267B1 co-purified with substrates (either flavanone or tetradecanoic acid) were thawed and concentrated further from ~20-35 mg/mL to ~60-70 mg/mL for

crystallization trials, as lower protein concentrations did not result in crystal formation. Sitting-drop vapor-diffusion method was used for the search of the crystallization growth conditions. The Mosquito crystallization robot (TTP LabTech) aided to prepare crystallization drops consisting of 100 nL : 100 nL, protein : reservoir ratio, that were set up against 50 µL of reservoir solution in 96-well MRC2 crystallization plates. The initial crystals grew from 35% (v/v) Tacsimate reagent (pH 7.0) after approximately 3 days of equilibration at 293 K. Crystals were then optimized manually by using the hanging-drop vapor diffusion method, performed in 24-well plates carrying 500 µL of reservoir solution. Optimized drops consisted of 2 µL protein and 2 µL reservoir solution. Single, red crystals grew within approximately 1 week of equilibration at 293 K, from 36% (v/v) Tacsimate reagent (pH 7.0).

Data collection and structure determination

Prior to data collection, the crystals were briefly passed through a drop of 70% (v/v) Tacsimate reagent (pH 7.0), serving as a cryo-protection solution and flash-cooled in the 110K-cold nitrogen gas cryostream. The data were collected at the European Synchrotron Radiation Facility (ESRF) in Grenoble, France (MX beamline ID29). The datasets were indexed and integrated using XDS [30], while scaling and merging was done with AIMLESS from the CCP4 software suite [31]. The structures were solved by the molecular replacement method and partially refined using the Auto-Rickshaw software pipeline [32], taking advantage of availability of the high-resolution data. The model of the P450 PikC (CYP107L1) D50N mutant (PDB entry 4UMZ, chain B) served as a search model (44 % sequence identity). Protein molecules in the crystals packed with 36% solvent content and the structures were solved in an orthorhombic space group ($P2_12_12_1$), containing 1 molecule per asymmetric unit. The initial, partially refined models of CYP267B1 were then further refined by performing iterative cycles of model building in COOT [33] and refinement with

Phenix.refine [34]. During last refinement rounds the waters and/or substrates were added to the models. Final model validation was carried out with Molprobity incorporated in the PHENIX software [35]. The final model coordinates and structure factor amplitudes were deposited in the Protein Data Bank (PDB): 6GK5 (CYP267B1) and 6GK6 (CYP267B1-MYR).

Molecular docking and structure analysis

Selected ligands were docked into the active site of CYP267B1 structure (receptor) with the use of Autodock Vina 1.1.2 [36]. The receptor model had all crystallographic waters removed prior to docking experiments. Coordinates for fatty acids and drug molecules were taken from the available crystal structures: decanoic acid (DKA, PDB entry 5IBO), dodecanoic acid (DAO, PDB entry 5UCA), tetradecanoic acid (MYR, PDB entry 4TKH), diclofenac (DIF, PDB entry 4UBS) and ibuprofen (IZP, PDB entry 3R8G). For flavanone, a SMILES string taken from PubChem (PubChem CID: 10251) was converted into PDB coordinates using the CACTUS web server (https://cactus.nci.nih.gov) and the ligand geometry was then optimized with PRODRG2 web server [37]. For all ligands, the correct number of rotatable bonds was confirmed by manual inspection in AutoDockTools 1.5.6. Hydrogens and Gasteiger charges were also added in the above-mentioned software. The protein was kept rigid during docking and the simulation cell consisted of a grid box, centered at the heme-iron, and covering the whole distal heme pocket (x:36 Å, y:46 Å, z:34 Å). Exhaustiveness parameter was default (= 8) for the fatty acids and set to 48 for the remaining three substrates. Binding poses were analyzed according to lowest binding energies and distances of the target oxidation sites in the ligand to the heme-iron.

Pairwise structural alignments were performed with PDBeFold [38]. Identification of substrate recognition sites (SRS) in CYP267B1 (UniProt: A9ERX9) was done by sequence

alignment with $P450_{cam}$ (UniProt: P00183) and according to description provided by Gotoh [39]: SRS1: 70-95 (B-C loop), SRS2: 171-177 (C' terminal part of F helix), SRS3: 183-194 (N' terminal part of G helix), SRS4: 231-249 (central part of I helix), SRS5: 286-296 (K-β2 connection) and SRS6: 390-397 (β3-hairpin). Substrate binding residues were analyzed in AutoDockTools 1.5.6 and LigPlot+ [40]. The molecular surface of the inner active site cavity was calculated by '3V: *Voss Volume Voxelator*' web server and using 1.4 Å inner probe radius [41]. All structural figures were prepared using UCSF Chimera [42]. QR codes in selected figures allow to inspect the image in 3D on a smartphone or tablet, by means of the Augment app [43].

Results/Discussion

Identification of new substrates for CYP267B1: binding and conversion

Even though CYP267B1 was already shown to bind and convert a variety of different substrates, it is of interest to further map its substrate scope and test new substrate classes for their potential to be converted. Of special interest are fatty acids (FAs) due to their biotechnological importance. Besides FAs, we also focused on flavonoids, which have a similar molecular mass as FAs but differ in chemical structure, and represent a large family of natural compounds with broad possibilities for application.

To prepare the enzyme for substrate-binding and conversion experiments, the CYP267B1 from *Sorangium cellulosum* So ce56 was recombinantly expressed in *E. coli* and purified by a two-step procedure involving affinity chromatography followed by size-exclusion chromatography, yielding about 300 nmol protein per liter of expression culture.

Protein purity was evaluated by SDS-PAGE (Supplementary Figure 1) and the purified enzyme showed its typical spectral properties as reported previously [26].

Next, it was tested whether the enzyme can accept FAs as substrates. The regioselective enzymatic oxidation catalyzed by P450s producing hydroxy-fatty acids (hydroxy-FAs) is a highly desired biocatalytic process, as hydroxy-FAs have many applications in the chemical, pharmaceutical, cosmetic or food industries [44,45]. Hydroxylated FAs are more stable and solvent miscible in comparison to their non-hydroxylated counterparts [46]. Use of hydroxy-FAs in the synthesis of resins and polymers allows to obtain more flexible and more resistant final products in comparison to those derived from petroleum [47]. FAs are natural substrates of many P450s and, most interestingly, previous bioinformatic analysis indicated that CYP267B1 clusters well with the known FA-hydroxylating P450 enzymes [21]. Moreover, the closely related CYP267A1 (43% sequence identity) was recently shown to convert medium- to long-chain saturated FAs [48]. Indeed, we found that CYP267B1 is able to bind three different medium-chain FAs, *i.e.*, decanoic acid (capric acid), dodecanoic acid (lauric acid) and tetradecanoic acid (myristic acid), as indicated by characteristic perturbations of the heme spectrum (a 'type I' shift of the Soret band from 418 nm towards 390 nm). Taking advantage of the substrate-binding induced spectral changes, dissociation constants (K_d) were determined for all three FAs (Table 1, Supplementary Figure 2), revealing that CYP267B1 has a relatively strong affinity for tetradecanoic acid (K_d = 2.1 ± 0.7μM) and weaker affinities for the two other (shorter) FAs (K_d = 44 ± 4.0 μM and 13 ± 1.0 μM for decanoic acid and dodecanoic acid, respectively). In comparison, CYP267A1 displays a very weak affinity for tetradecanoic acid (K_d could not be measured), but its affinity for decanoic acid and dodecanoic acid binding is stronger compared to CYP267B1 (K_d values of 2.97 ± 0.24 μM and 1.05 ± 0.01 μM, respectively) [48].

Next, flavanone (IUPAC: 2-phenyl-2,3-dihydrochromen-4-one) was investigated as a potential substrate of CYP267B1. This flavonoid compound is similar to tetradecanoic acid in molecular weight (224.3 g/mol versus 228.4 g/mol), but displays a completely different chemical structure (Supplementary Figure 3). Flavonoids are a large and diverse group of polyphenolic chemical compounds originating from plant sources [49]. They possess a three-ring structure usually decorated with different substitutions that promote their various antioxidant, anti-inflammatory, anti-allergic or anti-carcinogenic roles [50,51]. The flavanone structure is known to serve as precursor to different flavonoid structures so that it can be expected that its specifically oxidized P450-generated bioconversion products are valuable intermediates for the chemical or pharmaceutical industry. Like with the FAs, titrations of CYP267B1 with flavanone resulted in 'type I' heme spectral changes, indicating binding of the flavonoid (Supplementary Figure 4). The induced perturbations to the heme spectrum were much smaller than for the FAs (below 10% of the maximum spin shift), indicating that flavanone binding by CYP267B1 is weak, and a dissociation constant could not be measured.

After demonstrating their binding to CYP267B1, the three FAs and flavanone were subjected to *in vitro* conversion experiments. Truncated bovine adrenodoxin (Adx_{4-108}) and adrenodoxin reductase (AdR), shown before to support the functional activity of CYP267B1 as heterologous redox partner proteins [25], were applied to reconstitute the activity of the enzyme. All four tested substrates were converted *in vitro* by CYP267B1, with conversion ratios of 76%, 99%, 90% and 28% for decanoic acid, dodecanoic acid, tetradecanoic acid and flavanone, respectively. Taken together, it can be stated that two novel classes of CYP267B1 substrates were found, medium-chain FAs and a flavonoid compound. Subsequent investigations concentrated on the identification of formed products.

In-chain fatty acid hydroxylation by CYP267B1

GC-MS analysis of the product mixtures resulting from FA conversions by CYP267B1 revealed that the enzyme preferentially oxidized the inner carbon atoms of the FA chains and generated mono-hydroxylated products (Figure 1, Table 1). The three FAs are rather unselectively hydroxylated at their inner chain carbon atoms, with the highest regioselectivity exhibited towards decanoic acid (42% hydroxylation in ω-3 position). When using tetradecanoic acid, the formation of 28% ω-3 and 35% ω-4 hydroxylated metabolites was obtained (Table 1). The in-chain FA hydroxylations (occurring at positions between the methyl and carboxylate terminal carbons) are relatively common among bacterial P450s, usually resulting in a range of positions being hydroxylated during catalysis and multiple products being formed [52]. With increasing chain length of the FA substrate the regioselectivity of hydroxylation catalyzed by CYP267B1 is shifted from positions located closer to the methyl terminal carbon of the FA (ω-1, ω-2, ω-3) towards the inner-chain positions (ω-3 and ω-4). A relatively similar tendency to hydroxylate longer FAs at the further progressing inner-chain positions was previously reported for CYP276A1 from the same organism [48].

Regioselective hydroxylation of flavanone by CYP267B1

The *in vitro* conversion of flavanone by CYP267B1 resulted in one major product (P1, 15%) and some undefined side products (P2-P5) as detected by HPLC analysis (Figure 2A). In order to identify the products by NMR analysis, the previously established *E. coli*-based whole-cell system was used to scale-up the flavanone conversion [26]. In the applied *in vivo* system CYP267B1 was supported by an autologous redox partner pair, the FdR_B (Ferredoxin-$NADP^+$ reductase) and Fdx8 (Ferredoxin 8) from *Sorangium cellulosum* So ce56. As a result, about 40% of 200 μM flavanone was converted within 24 h (Figure 2B). The resulting product profile was similar to the one observed *in vitro*, with one major product (P1, 21%), one side product (P5) and three minor side products (P2, P3 and P4). For NMR

analysis only the P1 product was obtained in sufficient amounts (~10 mg), and by revealing a secondary hydroxyl group it was identified as 3-hydroxyflavanone (Supplementary Table 1). The H2 and H3 protons showed a coupling constant of 12.3 Hz leading to the conclusion that the product exhibits a *trans* configuration around the C2-C3 bond, in agreement with the literature data [53] (Figure 2C).

Biosynthesis of differently hydroxylated flavonoids is of high interest considering the difficulty of their extraction from plant sources or the necessity for complicated procedures to obtain them synthetically. Several papers have reported flavonoid hydroxylations catalyzed by P450s from mammals [54], plants [55], bacteria [56] or fungi [57]. To the best of our knowledge, production of 3-hydroxyflavanone with flavanone as substrate was only reported recently, in a whole-cell system expressing CYP110E1 from *Nostoc* sp. strain PCC 712; however, the final product co-existed with other side products and its final yields were not given [58]. Therefore, the myxobacterial CYP267B1 characterized in this work offers a new biocatalytic route to regioselective production of 3-hydroxyflavanone.

Overall structural features of CYP267B1

To gain insights into the structural basis of binding and regioselective conversion of flavanone and FAs by CYP267B1, the protein-substrate complexes were analyzed by X-ray crystallography. Protein samples were co-purified with flavanone or tetradecanoic acid and screened for suitable crystallization growth conditions. Tetradecanoic acid was selected for crystallographic binding analysis, considering that among the three tested FAs it contains the longest fatty acid chain and displayed the highest binding affinity. Crystals of CYP267B1 in the presence of flavanone or tetradecanoic acid grew at identical conditions using highly concentrated protein samples (~70 mg/mL). Crystal structures were determined by the molecular replacement method and refined at 1.6 Å resolution in the $P2_12_12_1$ space group,

with the asymmetric unit containing one protein molecule (see Table 2 for relevant crystallographic statistics). In both structures an ordered water molecule is occupying the sixth coordination position of the heme-iron. Additional clear density for a bound substrate was only visible in the case of tetradecanoic acid. In the structure obtained from a crystal grown in the presence of flavanone, the density for the flavanone molecule was very weak and the substrate could not be modeled. The two structures (CYP267B1 and CYP267B1-MYR) are highly similar, displaying a root mean square deviation [r.m.s.d] in Cα-backbone positions of 0.09 Å. The polypeptide chain is well resolved in the electron density maps, except for the first five N-terminal residues and most of the C-terminal His_6-tag, which appear disordered and were left out from the models.

The overall structure of CYP267B1 shows the characteristic fold for the P450 superfamily, with 16 α-helices and 3 β-sheets (1 two-stranded and 2 three-stranded β-sheets) (Figure 3). The heme iron is ligated to the conserved cysteine residue (Cys354) and the heme is further stabilized by several hydrophobic/van der Waals interactions with neighboring residues in the heme-binding pocket, and by ionic interactions of its propionates with the side chains of His99, Arg103, His352 and Arg296. The overall protein structure adopts a “closed” conformation, with the F and G helices and the F-G loop (loop connecting the helices F and G) topping the active site pocket and making it almost inaccessible to the solvent. According to a structural alignment performed against the whole PDB archive, the structure of CYP267B1 shows the highest match to “closed” crystal structures of PikC, another P450 accepting macrolides as substrates [59], (Figure 4A, Supplementary Table 2). Similar to PikC, the inner active site cavity of CYP267B1 is very spacious, consistent with its ability to accept a great variety of substrate sizes, ranging from very small ones like α-ionone [22] to large macrocyclic compounds such as epothilone D [24]. Towards the heme the pocket narrows, restricting access to the heme-iron, in particular due to the presence of residue

Leu92 located in the substrate recognition site 1 (SRS1) (Figure 4B). This residue corresponds to Phe87 in P450 BM3, and likely is similarly important for determining the substrate range and regioselectivity of hydroxylation [60]. Another residue in the SRS1 region of CYP267B1, His90, may further influence the regioselectivity of substrate hydroxylations. This is evident from comparing the structure of CYP267B1 with that of P450 EpoK, an enzyme from *Sorangium cellulosum* So ce90 involved in epothilone biosynthesis [61]. Both enzymes convert epothilone D, but a superposition with the structure of epothilone D-bound EpoK reveals that this substrate cannot adopt the same binding mode in the active site pocket of CYP267B1, due to a clash with His90 (Figure 4B). This difference may be related to previous observations that these enzymes convert epothilone D with quite different regioselectivities: while EpoK converts epothilone D exclusively to epothilone B, CYP267B1 oxidizes epothilone D to 5 different products [24]. The large pocket of CYP267B1 likely allows epothilone D to adopt several alternate binding modes resulting in the broad regioselectivity. The current crystal structure may help future engineering of CYP267B1 towards improved regioselective conversion of epothilone D, similar as done recently for P450 BM3 and the macrocylic substrate β-cembrenediol [62].

CYP267B1-tetradecanoic acid complex and prediction of fatty acid binding modes

As mentioned above, the extra electron density in the CYP267B1-MYR crystal structure allowed us to model a bound substrate (Figure 5A). The electron density displayed a continuous U-shape, consistent with a curved binding mode of the fatty acid. The ends of the electron density were not very clearly defined, though, therefore it was considered that MYR did bind in two alternative binding orientations (each modeled with 50% occupancy), either with its carboxylate end towards SRS1 (purple conformer, Figure 5B) and stabilized by a

water-mediated hydrogen bond with His90, or towards SRS5/SRS6 (pink conformer, Figure 5B), forming several water-mediated interactions with the backbone atoms of Glu291, Leu292, Ser293, Pro393 and Thr394. In both binding modes, MYR forms hydrophobic interactions with Leu92 (SRS1); Leu176 (SRS2); Val242, Ala243, Thr247 (SRS4); Ala290, Leu292 (SRS5) and Pro393 (SRS6) (Figure 5B). The orientation of MYR relative to the heme-iron points to predominant oxidation at the C7 or C8 carbon atoms of tetradecanoic acid (3.9-4.2 Å distance to the heme-iron in both binding modes), suggesting ω-6/ ω-7 hydroxylation, which is not observed experimentally. Thus, the alternate binding modes of MYR, while consistent with the electron density, are non-productive. It is not uncommon for FAs to be bound in non-productive conformations in P450 crystal structures: this has been observed previously in the palmitoleic acid-bound P450 BM3 structure [63] or the lauric acid-bound structure of CYP107L2 from *Streptomyces avermitilis* [64]. Moreover, the presence of the heme-bound water in the CYP267B1-MYR structure further indicates that the enzyme adopts a 'resting' state (low-spin heme-Fe); in order to reach the 'active' state during catalysis the substrate needs to displace this water molecule.

To predict plausible 'active' binding conformations of the three FA substrates identified in this work, docking calculations were performed with the crystal structure of CYP267B1. For each FA twenty binding poses were generated and ranked according to their lowest predicted binding energies. It was found that in great majority (14-16 out of 20 binding poses) FAs were predicted to bind in the top parts of the active site (between the central part of the B-C loop and the F helix/F-G loop). Only about 4-6 poses placed the individual FA relatively close to the heme, and among those only 3-4 poses had one or more carbon atoms at a distance suitable for hydroxylation. From those latter FA binding poses, the ones with the lowest predicted binding energies were selected for further analysis. Decanoic acid is predicted to bind with its carboxylate group in hydrogen bonding distance to His90

(SRS1) and interact with several active site residues like Leu92 (SRS1); Thr247, Ala243, Leu239, Val242 (SRS4); Ala290 (SRS5); Ile395 (SRS6). Its omega end is curved over the heme exposing the C7, C8, C9 carbon atoms for hydroxylation (Figure 6A). A very similar binding pose was obtained when dodecanoic acid was docked into the CYP267B1 active site, although its carboxylate group is too far from His90 (SRS1) to form a hydrogen bond (3.6 Å). Dodecanoic acid is predicted to interact with a similar set of residues: Leu92 (SRS1); Leu176 (SRS2); Phe238, Leu239, Val242, Ala243, Thr247 (SRS4); Ala290, Pro289 (SRS5) and Ile395 (SRS6). In such a predicted binding conformation the C7, C8, C9, C10 and C11 carbon atoms are most optimally located for hydroxylation (Figure 6B). In contrast to decanoic and dodecanoic acid binding, tetradecanoic acid docks in a distinct flipped orientation towards the SRS5/SRS6 residues and its pose resembles one of the alternative conformers found in the CYP267B1-MYR crystal structure. No hydrogen bonds are predicted to stabilize the bound MYR and the long molecule forms several hydrophobic interactions with residues in majority belonging to SRS4 and SRS5: Leu92 (SRS1); Ala243, Val242, Thr247, Leu239 (SRS4); Thr294, Leu292, Ser293, Ile295 (SRS5) and Pro393, Ile395 (SRS6). Importantly, the predicted MYR binding mode positions carbon atoms C9, C10, C11 and C12 at a close distance from the heme-iron (Figure 6C), consistent with its observed hydroxylation pattern.

Thus, our results demonstrate that the closed protein conformation displayed by the CYP267B1 crystal structure allows prediction of plausible ‘productive’ binding poses for each of the three FAs. The FAs bind with their omega-end bended over the heme, positioning different carbon atoms near to the heme-iron. The observed hydroxylation patterns are consistent with a very spacious active site pocket and lack of strong hydrogen bonding interactions, making it difficult for CYP267B1 to strictly control the conformation and position of the flexible FA chain within its active site. In contrast, terminal FA hydroxylases

(ω-hydroxylases) like CYP153A from *Marinobacter aquaeolei* [65], which are highly regioselective, display a narrow hydrophobic tunnel for binding FAs, restricting substrate motility in their active site. Likely, all in-chain FA hydroxylating P450s display similar active site properties as CYP267B1, explaining why so far no 'exclusive P450 in-chain hydroxylase' has been reported [52].

Prediction of flavanone binding mode

Using similar procedures also flavanone was docked into the active site of CYP267B1. Flavonone, like the docked FAs, preferred binding in the top parts of the active site (13 out of 20 poses). Nevertheless, among the binding poses close enough to the heme for hydroxylation the one with the lowest predicted binding energy was consistent with the major oxidation site at carbon C3 (Figure 6D). This productive binding pose of flavanone is stabilized by hydrogen bonding of its C4-carbonyl group to the backbone amide nitrogen of Ala243 (SRS4) and by hydrophobic interactions with the side chains of Leu176 (SRS2); Phe238, Leu239, Val242, Thr247 (SRS4); Ala290, Thr294 (SRS5) and Ile395 (SRS6). Importantly, the C3 carbon is oriented towards the heme-iron such that hydroxylation would lead to a *trans*-configuration around the C2-C3 bond, which is perfectly in agreement with our experimental data (Figure 6D).

Molecular docking of diclofenac and ibuprofen

Considering the clear results obtained from the docking calculations with flavanone and the three FAs, we were wondering whether the crystal structure of CYP267B1 also allowed plausible productive binding modes to be predicted for other known CYP267B1 substrates with a characterized product profile. Thus, two drug molecules, diclofenac and ibuprofen, were docked into the active site of CYP267B1, and for both substrates low energy binding

poses were identified (Figure 7) which showed good agreement with the previously identified protein activity for formation of 4’-hydroxydiclofenac and 2-hydroxyibuprofen as main products, respectively [26]. Diclofenac is predicted to expose the C4’ carbon for oxidation and to interact with Leu92 (SRS1); Leu176 (SRS2); Leu292, Ser293, Thr294 (SRS5) and Pro393, Thr394, Ile395 (SRS6) (Figure 7A). The ibuprofen molecule is predicted to form one hydrogen bond with His90 (SRS1) and be further stabilized by Leu92 (SRS1); Leu176 (SRS2); Phe238, Leu239, Val242, Ala243 (SRS4); Ala290 (SRS5) and Ile395 (SRS6) (Figure 7B).

In conclusion, the present study identified two novel substrate classes to be converted by the broad substrate range enzyme, cytochrome P450 CYP267B1 from *Sorangium cellulosum* So ce56. It was shown to oxidize medium-chain fatty acids and flavanone. For the first time a three-dimensional structure of this versatile enzyme was determined, allowing to dock several substrates identified in this work and before, and to predict plausible productive conformations. However, in order to predict the reaction products more precisely in the future (*e.g.* from substrates with unidentified product profiles), more sophisticated bioinformatic methods could be employed like molecular dynamics (MD) simulations or combined quantum mechanical/molecular mechanical (QM/MM) calculations (also involving the heme oxoferryl moiety). Thus, the structure of CYP267B1 will likely inspire future rational mutagenesis-based adaptations of this protein towards different biotechnological applications.

Acknowledgements

We thank Dr. Josef Zapp for measurment of the NMR samples, Birgit Heider-Lips for purification of $Adx_{4\text{-}108}$ and AdR, and ID29 beamline staff of the European Synchrotron Radiation Facility (ESRF) for assistance during remote data collection.

Declarations of interest

The authors declare no competing interests.

Funding information

The work was supported by a grant from the Deutsche Forschungsgemeinschaft to RB (Be1343/23). IKJ was funded by the People Programme (Marie Curie Actions) of the European Union's 7th Framework Programme (FP7/2007-2013) under REA Grant Agreement 289217 (ITN P4FIFTY) grant to the University of Groningen.

Author contribution

YK conceived the study. ML, YK, AS, MG performed all biochemical experiments and collected data. ML, YK, AS, MG, VU, RB analyzed and interpreted the biochemical data. IKJ crystallized the protein, collected crystallographic data, determined the crystal structures and performed docking studies. IKJ and AMWHT analyzed and interpreted the structural data. IKJ wrote the paper. All authors contributed to writing of the manuscript, and read and approved the final version of the paper.

References

1 Nelson, D. R. (2017) Cytochrome P450 diversity in the tree of life. Biochim. Biophys. Acta. **1866**, 141-154

2 Munro, A. W., Girvan, H. M., Mason, A. E., Dunford, A. J. and McLean, K. J. (2013) What makes a P450 tick? Trends Biochem. Sci. **38**, 140-150

3 Mizutani, M. and Sato, F. (2011) Unusual P450 reactions in plant secondary metabolism. Arch. Biochem. Biophys. **507**, 194-203

4 Sono, M., Roach, M. P., Coulter, E. D. and Dawson, J. H. (1996) Heme-containing oxygenases. Chem. Rev. **96**, 2841-2888

5 Denisov, I. G., Makris, T. M., Sligar, S. G. and Schlichting, I. (2005) Structure and chemistry of cytochrome P450. Chem. Rev. **105**, 2253-2277

6 Hannemann, F., Bichet, A., Ewen, K. M. and Bernhardt, R. (2007) Cytochrome P450 systems—biological variations of electron transport chains. Biochim. Biophys. Acta. **1770**, 330-344

7 Girvan, H. M. and Munro, A. W. (2016) Applications of microbial cytochrome P450 enzymes in biotechnology and synthetic biology. Curr. Opin. Chem. Biol. **31**, 136-145

8 Urlacher, V. B. and Girhard, M. (2012) Cytochrome P450 monooxygenases: An update on perspectives for synthetic application. Trends Biotechnol. **30**, 26-36

9 Bernhardt, R. (2006) Cytochromes P450 as versatile biocatalysts. J. Biotechnol. **124**, 128-145

10 Bernhardt, R. and Urlacher, V. B. (2014) Cytochromes P450 as promising catalysts for biotechnological application: Chances and limitations. Appl. Microbiol. Biotechnol. **98**, 6185-6203

11 von Buhler, C., Le-Huu, P. and Urlacher, V. B. (2013) Cluster screening: An effective approach for probing the substrate space of uncharacterized cytochrome P450s. Chembiochem. **14**, 2189-2198

12 Khatri, Y., Hannemann, F., Girhard, M., Kappl, R., Meme, A., Ringle, M., Janocha, S., Leize-Wagner, E., Urlacher, V. B. and Bernhardt, R. (2013) Novel family members of CYP109 from *Sorangium cellulosum* So ce56 exhibit characteristic biochemical and biophysical properties. Biotechnol. Appl. Biochem. **60**, 18-29

13 Putkaradze, N., Litzenburger, M., Abdulmughni, A., Milhim, M., Brill, E., Hannemann, F. and Bernhardt, R. (2017) CYP109E1 is a novel versatile statin and terpene oxidase from *Bacillus megaterium*. Appl. Microbiol. Biotechnol. **101**, 8379-8393

14 Yin, Y. C., Yu, H. L., Luan, Z. J., Li, R. J., Ouyang, P. F., Liu, J. and Xu, J. H. (2014) Unusually broad substrate profile of self-sufficient cytochrome P450 monooxygenase CYP116B4 from *Labrenzia aggregata*. Chembiochem. **15**, 2443-2449

15 Acevedo-Rocha, C., Gamble, C., Lonsdale, R., Li, A., Nett, N., Hoebenreich, S., Lingnau, J. B., Wirtz, C., Fares, C., Hinrichs, H., Deege, A., Mulholland, A. J., Nov, Y., Leys, D., McLean, K. J., Munro, A. W. and Reetz, M. T. (2018) P450-catalyzed regio- and diastereoselective steroid hydroxylation: Efficient directed evolution enabled by mutability landscaping. ACS Catal. **8**, 3395-3410

16 Weber, E., Seifert, A., Antonovici, M., Geinitz, C., Pleiss, J. and Urlacher, V. B. (2011) Screening of a minimal enriched P450 BM3 mutant library for hydroxylation of cyclic and acyclic alkanes. Chem. Commun. (Camb). **47**, 944-946

17 Lewis, J. C., Mantovani, S. M., Fu, Y., Snow, C. D., Komor, R. S., Wong, C. H. and Arnold, F. H. (2010) Combinatorial alanine substitution enables rapid optimization of cytochrome P450BM3 for selective hydroxylation of large substrates. Chembiochem. **11**, 2502-2505

18 Sowden, R. J., Yasmin, S., Rees, N. H., Bell, S. G. and Wong, L. L. (2005) Biotransformation of the sesquiterpene (+)-valencene by cytochrome P450cam and P450BM-3. Org. Biomol. Chem. **3**, 57-64

19 Whitehouse, C. J., Bell, S. G., Tufton, H. G., Kenny, R. J., Ogilvie, L. C. and Wong, L. L. (2008) Evolved CYP102A1 (P450BM3) variants oxidise a range of non-natural substrates and offer new selectivity options. Chem. Commun. (Camb). **28**, 966-968

20 Schneiker, S., Perlova, O., Kaiser, O., Gerth, K., Alici, A., Altmeyer, M. O., Bartels, D., Bekel, T., Beyer, S., Bode, E., Bode, H. B., Bolten, C. J., Choudhuri, J. V., Doss, S., Elnakady, Y. A., Frank, B., Gaigalat, L., Goesmann, A., Groeger, C., Gross, F., Jelsbak, L., Jelsbak, L., Kalinowski, J., Kegler, C., Knauber, T., Konietzny, S., Kopp, M., Krause, L., Krug, D., Linke, B., Mahmud, T., Martinez-Arias, R., McHardy, A. C., Merai, M., Meyer, F., Mormann, S., Munoz-Dorado, J., Perez, J., Pradella, S., Rachid, S., Raddatz, G., Rosenau, F., Ruckert, C., Sasse, F., Scharfe, M., Schuster, S. C., Suen, G., Treuner-Lange, A., Velicer, G. J., Vorholter, F. J., Weissman, K. J., Welch, R. D., Wenzel, S. C., Whitworth, D. E., Wilhelm, S., Wittmann, C., Blocker, H., Puhler, A. and Muller, R. (2007) Complete genome sequence of the myxobacterium *Sorangium cellulosum*. Nat. Biotechnol. **25**, 1281-1289

21 Khatri, Y., Hannemann, F., Ewen, K. M., Pistorius, D., Perlova, O., Kagawa, N., Brachmann, A. O., Muller, R. and Bernhardt, R. (2010) The CYPome of S*orangium cellulosum* So ce56 and identification of CYP109D1 as a new fatty acid hydroxylase. Chem. Biol. **17**, 1295-1305

22 Litzenburger, M. and Bernhardt, R. (2016) Selective oxidation of carotenoid-derived aroma compounds by CYP260B1 and CYP267B1 from *Sorangium cellulosum* So ce56. Appl. Microbiol. Biotechnol. **100**, 4447-4457

23 Rinkel, J., Litzenburger, M., Bernhardt, R. and Dickschat, J.S. (2018) An isotopic labelling strategy to study cytochrome P450 oxidations of terpenes. ChemBioChem. doi: 10.1002/cbic.201800215

24 Kern, F., Dier, T. K., Khatri, Y., Ewen, K. M., Jacquot, J. P., Volmer, D. A. and Bernhardt, R. (2015) Highly efficient CYP167A1 (EpoK) dependent epothilone B formation and production of 7-ketone epothilone D as a new epothilone derivative. Sci. Rep. **5**, 14881

25 Litzenburger, M., Kern, F., Khatri, Y. and Bernhardt, R. (2015) Conversions of tricyclic antidepressants and antipsychotics with selected P450s from *Sorangium cellulosum* So ce56. Drug Metab. Dispos. **43**, 392-399

26 Kern, F., Khatri, Y., Litzenburger, M. and Bernhardt, R. (2016) CYP267A1 and CYP267B1 from *Sorangium cellulosum* So ce56 are highly versatile drug metabolizers. Drug Metab. Dispos. **44**, 495-504

27 Uhlmann, H., Beckert, V., Schwarz, D. and Bernhardt, R. (1992) Expression of bovine adrenodoxin in *E. coli* and site-directed mutagenesis of /2 fe-2S/ cluster ligands. Biochem. Biophys. Res. Commun. **188**, 1131-1138

28 Sagara, Y., Wada, A., Takata, Y., Waterman, M. R., Sekimizu, K. and Horiuchi, T. (1993) Direct expression of adrenodoxin reductase in *Escherichia coli* and the functional characterization. Biol. Pharm. Bull. **16**, 627-630

29 Schenkman, J. B. and Jansson, I. (1998) Spectral analyses of cytochromes P450. Methods Mol. Biol. **107**, 25-33

30 Kabsch, W. (2010) XDS. Acta Crystallogr. D Biol. Crystallogr. **66**, 125-132

31 Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. **67**, 235-242

32 Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S. and Tucker, P. A. (2005) Auto-rickshaw: An automated crystal structure determination platform as an efficient tool for the validation of an X-ray diffraction experiment. Acta Crystallogr. D Biol. Crystallogr. **61**, 449-457

33 Emsley, P. and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. **60**, 2126-2132

34 Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. and Adams, P. D. (2012) Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. D Biol. Crystallogr. **68**, 352-367

35 Chen, V. B., Arendall, W. B.,3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. and Richardson, D. C. (2010) MolProbity: All-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. **66**, 12-21

36 Trott, O. and Olson, A. J. (2010) AutoDock vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comput. Chem. **31**, 455-461

37 Schuttelkopf, A. W. and van Aalten, D. M. (2004) PRODRG: A tool for high-throughput crystallography of protein-ligand complexes. Acta Crystallogr. D Biol. Crystallogr. **60**, 1355-1363

38 Krissinel, E. and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr. D Biol. Crystallogr. **60**, 2256-2268

39 Gotoh, O. (1992) Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. J. Biol. Chem. **267**, 83-90

40 Laskowski, R. A. and Swindells, M. B. (2011) LigPlot+: Multiple ligand-protein interaction diagrams for drug discovery. J. Chem. Inf. Model. **51**, 2778-2786

41 Voss, N. R. and Gerstein, M. (2010) 3V: Cavity, channel and cleft volume calculator and extractor. Nucleic Acids Res. **38**, W555-W562

42 Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. and Ferrin, T. E. (2004) UCSF chimera--a visualization system for exploratory research and analysis. J. Comput. Chem. **25**, 1605-1612

43 Wolle, P., Müller, M. P. and Rauh, D. (2018) Augmented reality in scientific Publications-Taking the visualization of 3D structures to the next level. ACS Chem. Biol. **13**, 496-499

44 Kim, K. and Oh, D. (2013) Production of hydroxy fatty acids by microbial fatty acid-hydroxylation enzymes. Biotechnol. Adv. **31**, 1473-1485

45 Ciaramella, A., Minerdi, D. and Gilardi, G. (2017) Catalytically self-sufficient cytochromes P450 for green production of fine chemicals. Rendiconti Lincei. **28**, 169-181

46 Metzger, J. O. and Bornscheuer, U. (2006) Lipids as renewable resources: Current state of chemical and biotechnological conversion and diversification. Appl. Microbiol. Biotechnol. **71**, 13-22

47 Burdock, G. A., Carabin, I. G. and Griffiths, J. C. (2006) Toxicology and pharmacology of sodium ricinoleate. Food Chem. Toxicol. **44**, 1689-1698

48 Khatri, Y., Hannemann, F., Girhard, M., Kappl, R., Hutter, M., Urlacher, V. B. and Bernhardt, R. (2015) A natural heme-signature variant of CYP267A1 from *Sorangium cellulosum* So ce56 executes diverse omega-hydroxylation. FEBS J. **282**, 74-88

49 Panche, A. N., Diwan, A. D. and Chandra, S. R. (2016) Flavonoids: An overview. J Nutr. Sci. **5**, e47

50 Heim, K. E., Tagliaferro, A. R. and Bobilya, D. J. (2002) Flavonoid antioxidants: Chemistry, metabolism and structure-activity relationships. J Nutr. Biochem. **13**, 572-584

51 Middleton, E., Kandaswami, C. and Theoharides, T. C. (2000) The effects of plant flavonoids on mammalian cells:Implications for inflammation, heart disease, and cancer. Pharmacol. Rev. **52**, 673-751

52 Hammerer, L., Winkler, C. K. and Kroutil, W. (2017) Regioselective biocatalytic hydroxylation of fatty acids by cytochrome P450s. Catalysis Letters. **148**, 787-812

53 Janzsó, G., Kállay, F., Koczor, I. and Radics, L. (1967) Stereochemistry of 3-hydroxy- and 3-acetoxyflavanone oximes. Tetrahedron. **23**, 3699-3704

54 Kakimoto, K., Murayama, N., Takenaka, S., Nagayoshi, H., Lim, Y. R., Kim, V., Kim, D., Yamazaki, H., Komori, M., Guengerich, F. P. and Shimada, T. (2018) Cytochrome P450 2A6 and other human P450 enzymes in the oxidation of flavone and flavanone. Xenobiotica. **29**, 1-12

55 Ayabe, S. and Akashi, T. (2006) Cytochrome P450s in flavonoid metabolism. Phytochem. Rev. **5**, 271-282

56 Liu, L., Yao, Q., Ma, Z., Ikeda, H., Fushinobu, S. and Xu, L. (2016) Hydroxylation of flavanones by cytochrome P450 105D7 from *Streptomyces avermitilis*. J Mol. Catal. B: Enzym. **132**, 91-97

57 Kasai, N., Ikushiro, S., Hirosue, S., Arisawa, A., Ichinose, H., Wariishi, H., Ohta, M. and Sakaki, T. (2009) Enzymatic properties of cytochrome P450 catalyzing 3′-hydroxylation of naringenin from the white-rot fungus *Phanerochaete chrysosporium*. Biochem. Biophys. Res. Commun. **387**, 103-108

58 Makino, T., Otomatsu, T., Shindo, K., Kitamura, E., Sandmann, G., Harada, H. and Misawa, N. (2012) Biocatalytic synthesis of flavones and hydroxyl-small molecules by recombinant *Escherichia coli* cells expressing the cyanobacterial CYP110E1 gene. Microbial Cell Factories. **11**, 95

59 Sherman, D. H., Li, S., Yermalitskaya, L. V., Kim, Y., Smith, J. A., Waterman, M. R. and Podust, L. M. (2006) The structural basis for substrate anchoring, active site selectivity, and product formation by P450 PikC from *Streptomyces venezuelae*. J. Biol. Chem. **281**, 26289-26297

60 Seifert, A., Vomund, S., Grohmann, K., Kriening, S., Urlacher, V. B., Laschat, S. and Pleiss, J. (2009) Rational design of a minimal and highly enriched CYP102A1 mutant library with improved regio-, stereo- and chemoselectivity. Chembiochem. **10**, 853-861

61 Nagano, S., Li, H., Shimizu, H., Nishida, C., Ogura, H., Ortiz de Montellano, P. R. and Poulos, T. L. (2003) Crystal structures of epothilone D-bound, epothilone B-bound, and substrate-free forms of cytochrome P450epoK. J. Biol. Chem. **278**, 44886-44893

62 Petrović, D., Bokel, A., Allan, M., Urlacher, V. B. and Strodel, B. (2018) Simulation-guided design of cytochrome P450 for chemo- and regioselective macrocyclic oxidation. J. Chem. Inf. Model. doi: 10.1021/acs.jcim.8b00043

63 Li, H. and Poulos, T. L. (1997) The structure of the cytochrome p450BM-3 haem domain complexed with the fatty acid substrate, palmitoleic acid. Nat. Struct. Biol. **4**, 140-146

64 Han, S., Pham, T., Kim, J., Lim, Y., Park, H., Jeong, D., Yun, C., Chun, Y., Kang, L. and Kim, D. (2017) Structural insights into the binding of lauric acid to CYP107L2 from *Streptomyces avermitilis.* Biochem. Biophys. Res. Commun. **482**, 902-908

65 Hoffmann, S. M., Danesh-Azari, H., Spandolf, C., Weissenborn, M. J., Grogan, G. and Hauer, B. (2016) Structure-guided redesign of CYP153AM.aq for the improved terminal hydroxylation of fatty acids. ChemCatChem. **8**, 3178-3178

66 Peterson JA, G. S. (1998) A close family resemblance: The importance of structure in understanding cytochromes P450. Structure **6**, 1079-1085

Tables

Table 1: Substrate binding affinity and product distribution of saturated fatty acid conversion by CYP267B1.

Substrate	K_d [µM]	Regioselectivity [%][a]						
		ω-7	ω-6	ω-5	ω-4	ω-3	ω-2	ω-1
Decanoic acid (C10:0)	44 ± 4.0	-	-	-	5	42	19	17
Dodecanoic acid (C12:0)	13 ± 1.0	-	<0.5	13	25	36	20	5
Tetradecanoic acid (C14:0)	2.1 ± 0.7	1	3	14	35	28	7	1

Note: [a] values are given as percent of the total product observed (mean of two individual experiments). -, not observed

Table 2: Crystallographic data collection and refinement statistics of CYP267B1.

	CYP267B1	**CYP267B1-MYR**
PDB code	**6GK5**	**6GK6**
Model statistics		
Monomers in the AU	1	1
Solvent content (%)	36	37
Ligands	n/a	myristic acid (MYR)
Data collection		
Beamline (ESRF)	ID29	ID29
Wavelength (Å)	0.97625	0.97625
Resolution range (Å)	41-1.60 (1.63-1.60)[a]	41-1.60 (1.63-1.60)
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters		
a, b, c (Å)	51.50 66.39 104.30	51.80 66.65 104.52
α, β, γ °	90 90 90	90 90 90
Observed reflections	175,537 (8,609)	169,339 (8,117)
Unique reflections	45,694 (2,319)	47,055 (2,350)
Multiplicity	3.8 (3.7)	3.6 (3.5)
Completeness (%)	95.6 (95.1)	97.5 (99.0)
$\langle I/\sigma(I)\rangle$	18.6 (3.2)	10.4 (2.2)
R_{merge} (%)	3.0 (32.5)	5.5 (48.5)
$R_{p.i.m.}$	2.3 (26.9)	4.1 (37.9)
$CC_{1/2}$ (%)	99.9 (83.7)	99.8 (67.6)
Refinement		
R_{work} (%)	18.7	18.7
R_{free} (%)	20.8	20.8
R.m.s.deviation, bond lengths (Å)	0.019	0.015
R.m.s.deviation, bond angles (°)	1.524	1.329
Average B-factors (Å)^2		
Overall	31.0	31.4
Protein	31.2	31.3
Heme	20.8	20.9
Myristic acid (MYR)	n/a	42.5
Ramachandran plot statistics		
Most favored (%)	96.8	97.0
Allowed regions (%)	2.9	3.0
Disallowed regions (%)	0.3	0
Molprobity overall score	1.31	1.36

[a] – Values in parentheses refer to the highest resoultion shell.

Figure legends

Figure 1: Gas chromatogram of ω-hydroxylation of decanoic acid (top), dodecanoic acid (middle) and tetradecanoic acid (bottom) catalyzed *in vitro* by CYP267B1. The '*' represents impurities during GC-MS measurement. The insets show fatty acid chemical structures with green arrows indicating hydroxylation positions (with major ones shown in bold).

Figure 2: Conversions of flavanone by CYP267B1. (A) HPLC chromatogram depicting the result of the CYP267B1-catalyzed *in vitro* conversion of flavanone. (B) HPLC chromatogram showing the result of the CYP267B1-catalyzed *in vivo* conversion of flavanone. 'Sub' designates the corresponding substrate. The main product (P1, t_R= 24.2 min) was isolated, purified and analyzed via NMR spectroscopy. (C) Scheme showing the CYP267B1 catalyzed reaction of flavanone hydroxylation. The relative configuration (*rel*) of the C2-phenyl and C3-hydroxyl groups was determined as *trans*, based on the coupling constant of 12.3 Hz between the H2 and H3 protons.

Figure 3: Overall fold of CYP267B1 from *Sorangium cellulosum* So ce56, colored from blue at the N' terminus to red at the C' terminus. The heme is shown as a red stick model in the center of the molecule. Helices are labeled A to L and three β-sheets are indicated, following the common P450 nomenclature [66]. QR code allows to appreciate the figure in 3D [43].

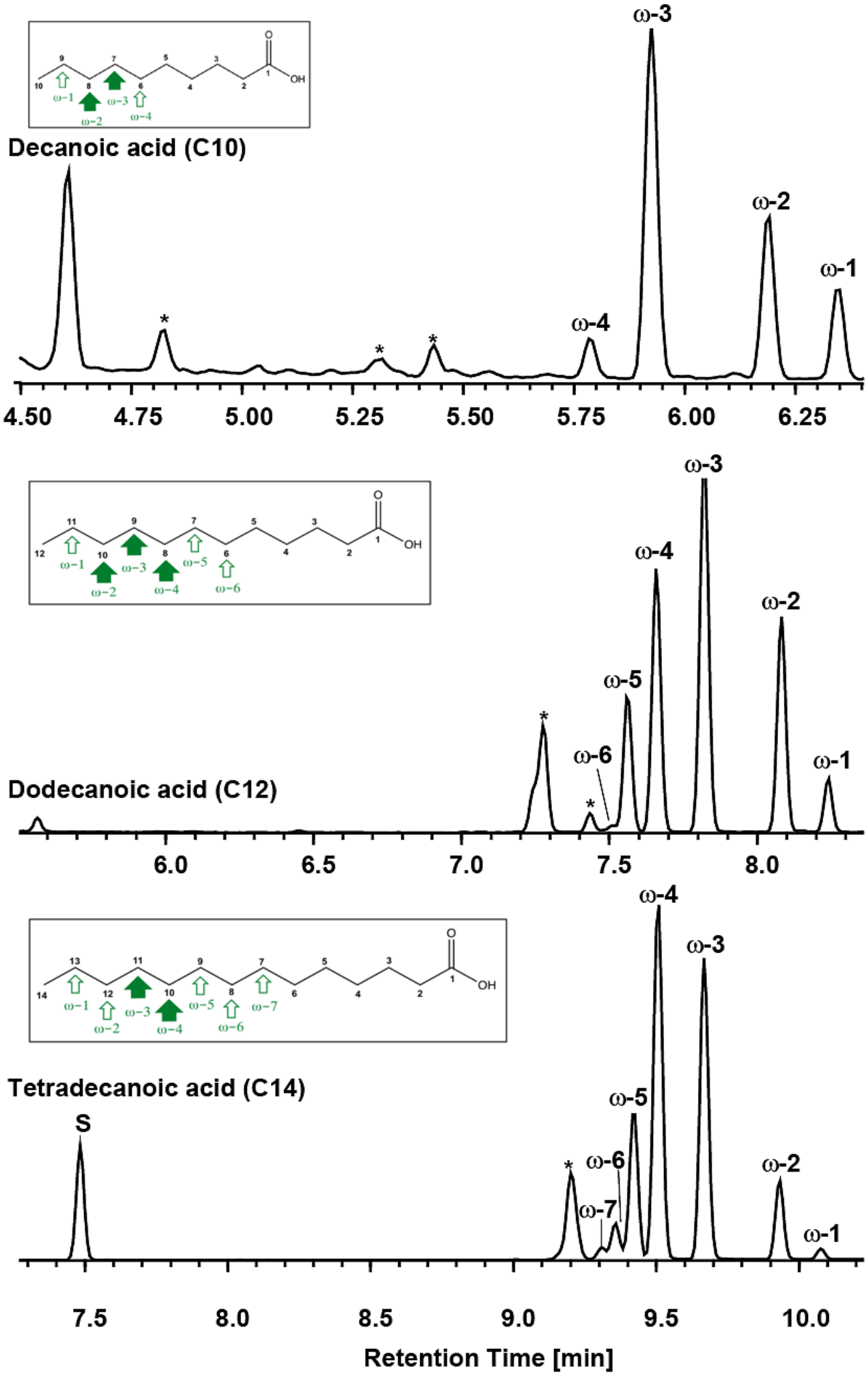
Figure 4: (A) Structural superposition of CYP267B1 in dark blue, with P450 PikC presenting a closed conformation (molecule A of PDB entry 2BVJ) in light blue, and PikC in an open conformation (molecule B of PDB entry 2BVJ) in magenta. (B) View of the inner active site cavity in CYP276B1. The solvent-accessible surface is shown as a semitransparent grey surface. A molecule of epothilone D as bound in the active site of P450 EpoK (PDB entry 1PKF) is superimposed for comparison and shows the steric hindrance caused by His90 of CYP267B1. The heme is depicted as red stick model. QR codes allow to appreciate the figure in 3D [43].

Figure 5: Tetradecanoic acid binding in CYP267B1. (A) Quality of electron density for the active site ligands in CYP267B1-MYR is depicted. Green mesh represents a composite $2F_o$ - F_c omit map calculated at 1.6 Å resolution and contoured at 0.5 σ. Black mesh is the final $2F_o$ - F_c electron density map, contoured at 0.5 σ. (B) Two alternative binding orientations of

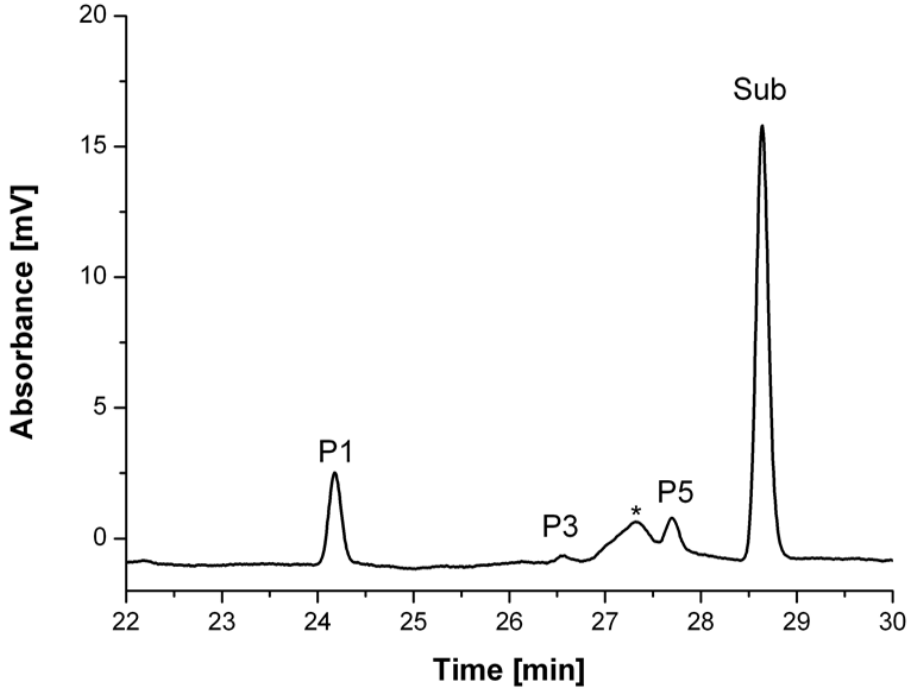
tetradecanoic acid and its binding interactions with active site waters and protein residues. One of the conformers is shown in purple with carbon atoms numbered in black, whereas the other is depicted in pink and its carbon atoms are numbered in red. QR code allows to appreciate the figure in 3D [43].

Figure 6: Molecular docking of fatty acids and flavanone in the active site of CYP267B1. (A) Decanoic acid is shown as an orange ball-and-stick model. (B) Dodecanoic acid is shown as ball-and-stick model, in green. (C) Tetradecanoic acid is depicted as ball-and-stick model in magenta. Carbon atoms are numbered in black. (D) Molecular model of flavanone binding in the active site of CYP267B1. The binding pose found closest to the heme and having lowest binding energy is shown, depicting a suitable conformation of flavanone (in cyan) for major hydroxylation event at carbon C3. Hydrogens are shown as white balls. The binding pose suggests generation of hydroxylation product being in *trans* configuration with the C2-phenyl group, what is in good agreement with the experimental data. Heme is shown as red stick model. QR codes allow to appreciate the figure in 3D [43].

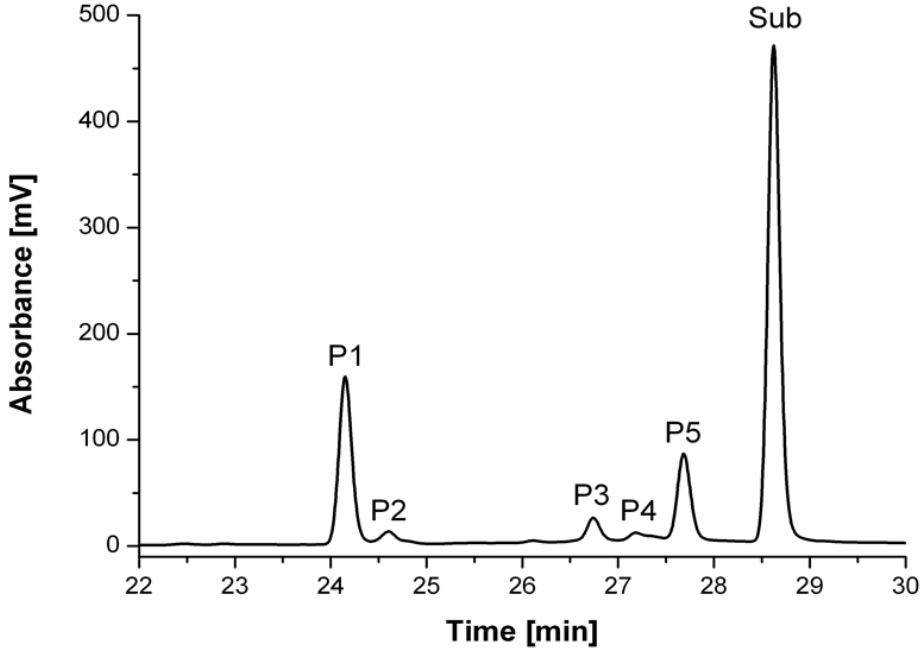
Figure 7: Molecular models of diclofenac and ibuprofen binding in the CYP267B1 active site. (A) Diclofenac is shown as light pink ball-and-stick model. (B) Ibuprofen is depicted as light blue ball-and-stick model. Both binding modes are consistent with the available literature data reporting generation of 4'-OH-diclofenac and 2-OH-ibuprofen by CYP267B1 from *Sorangium cellulosum* So ce56. Heme is shown as red stick model. QR codes allow to appreciate the figure in 3D [43].


Decanoic acid (C10)
ω-1
ω-2
ω-3
ω-4
4.50
4.75
5.00
5.25
5.50
5.75
6.00
6.25
Dodecanoic acid (C12)
ω-1
ω-2
ω-3
ω-4
ω-5
ω-6
6.0
6.5
7.0
7.5
8.0
Tetradecanoic acid (C14)
S
ω-1
ω-2
ω-3
ω-4
ω-5
ω-6
ω-7
7.5
8.0
8.5
9.0
9.5
10.0
Retention Time [min]

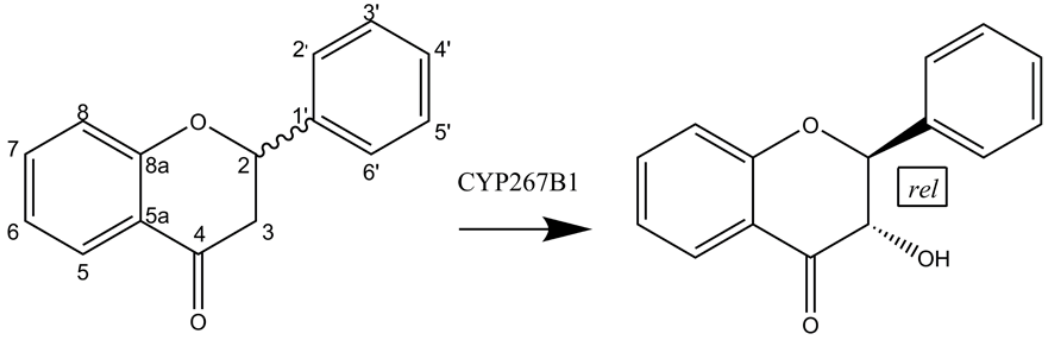
A

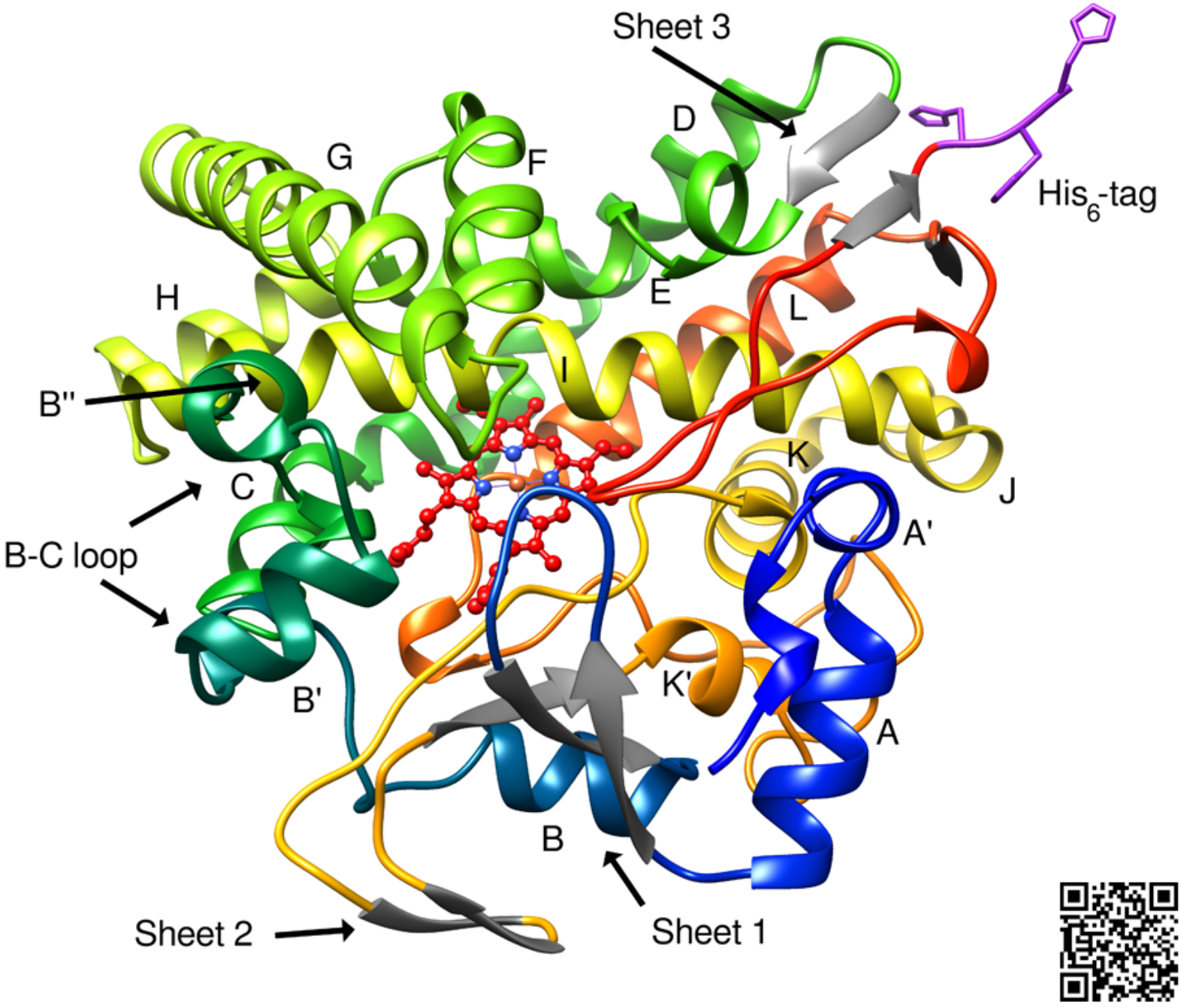

Absorbance [mV]
Time [min]
Sub
P1
P3
P5
*

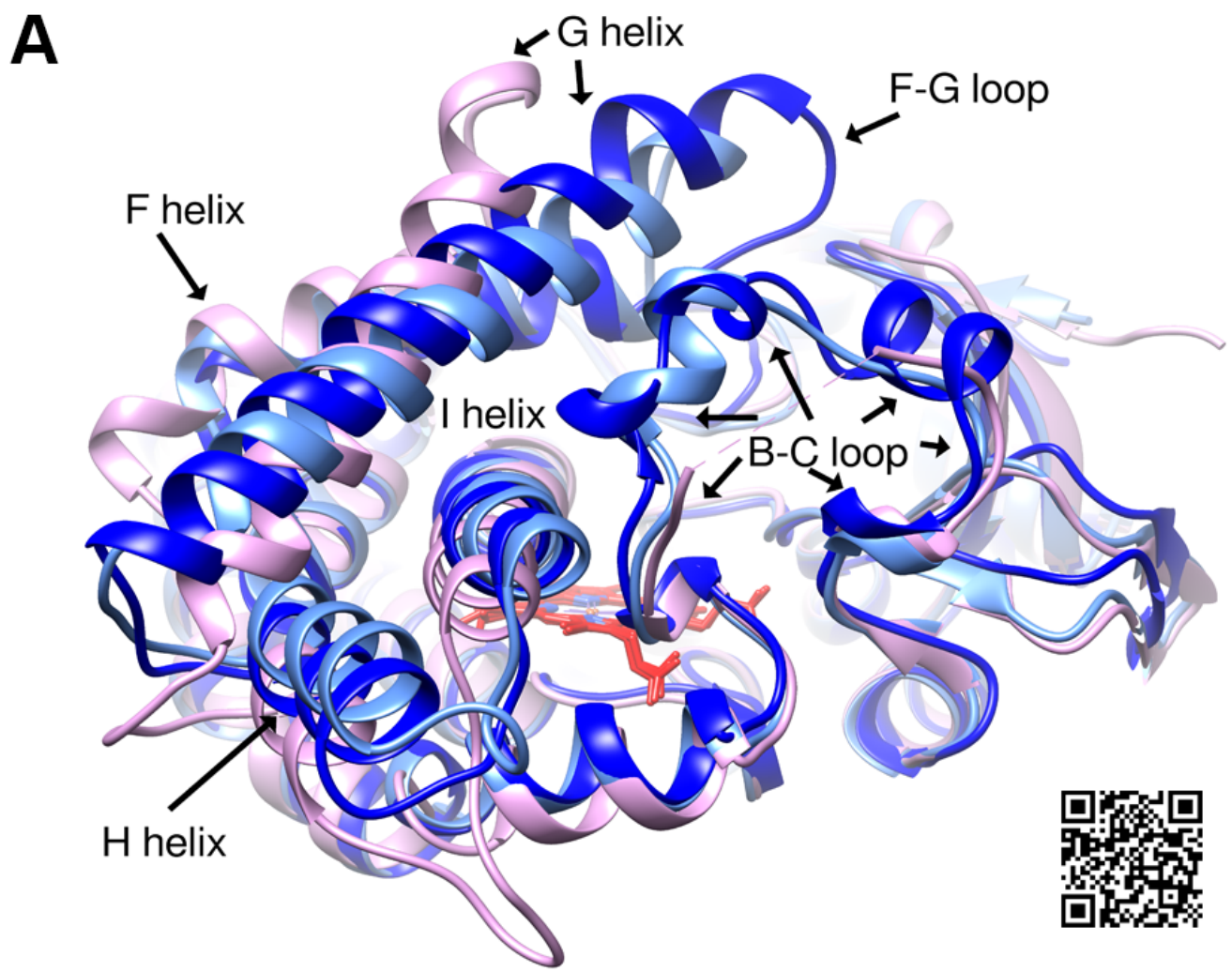
B

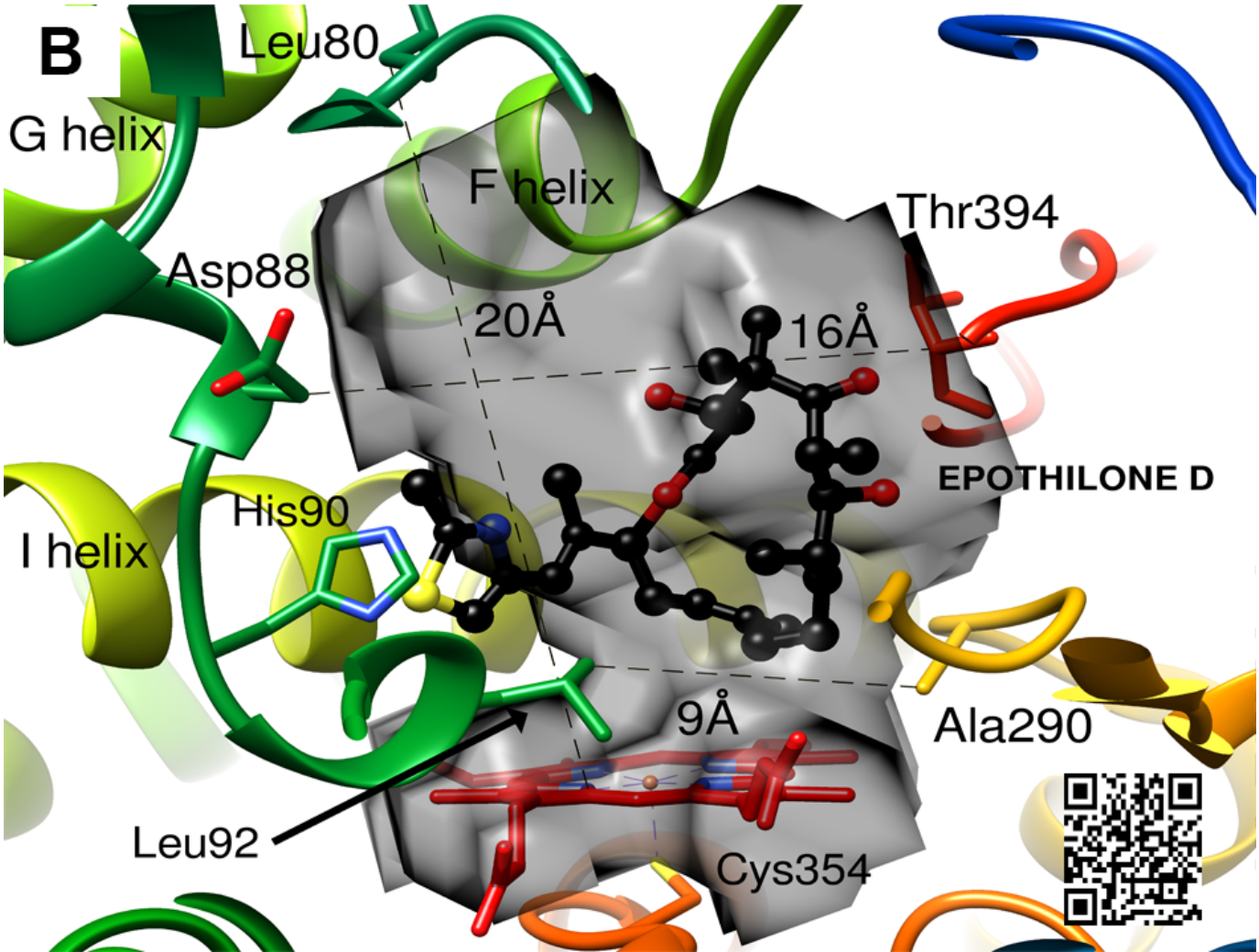

Absorbance [mV]
Time [min]
Sub
P1
P2
P3
P4
P5

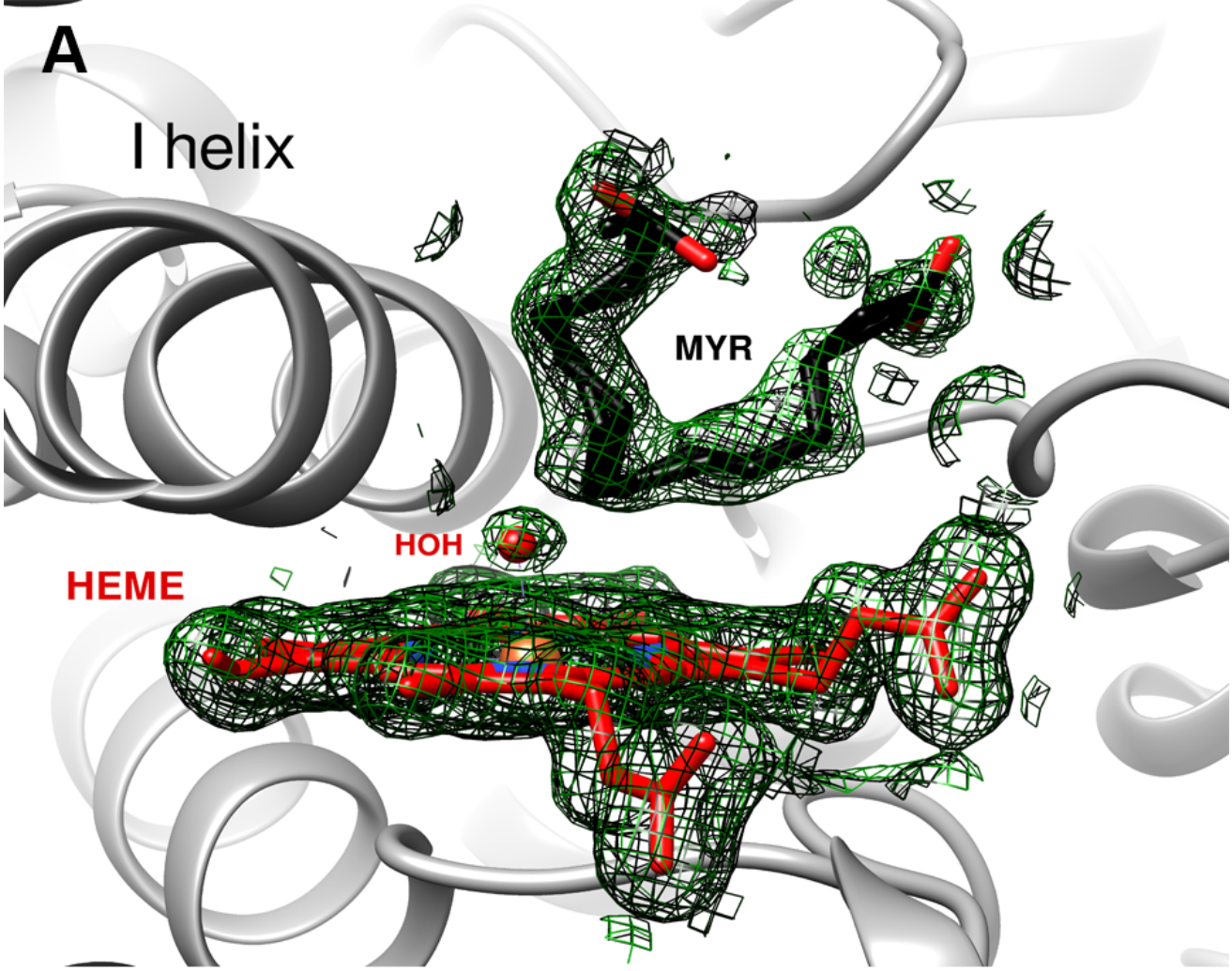
C

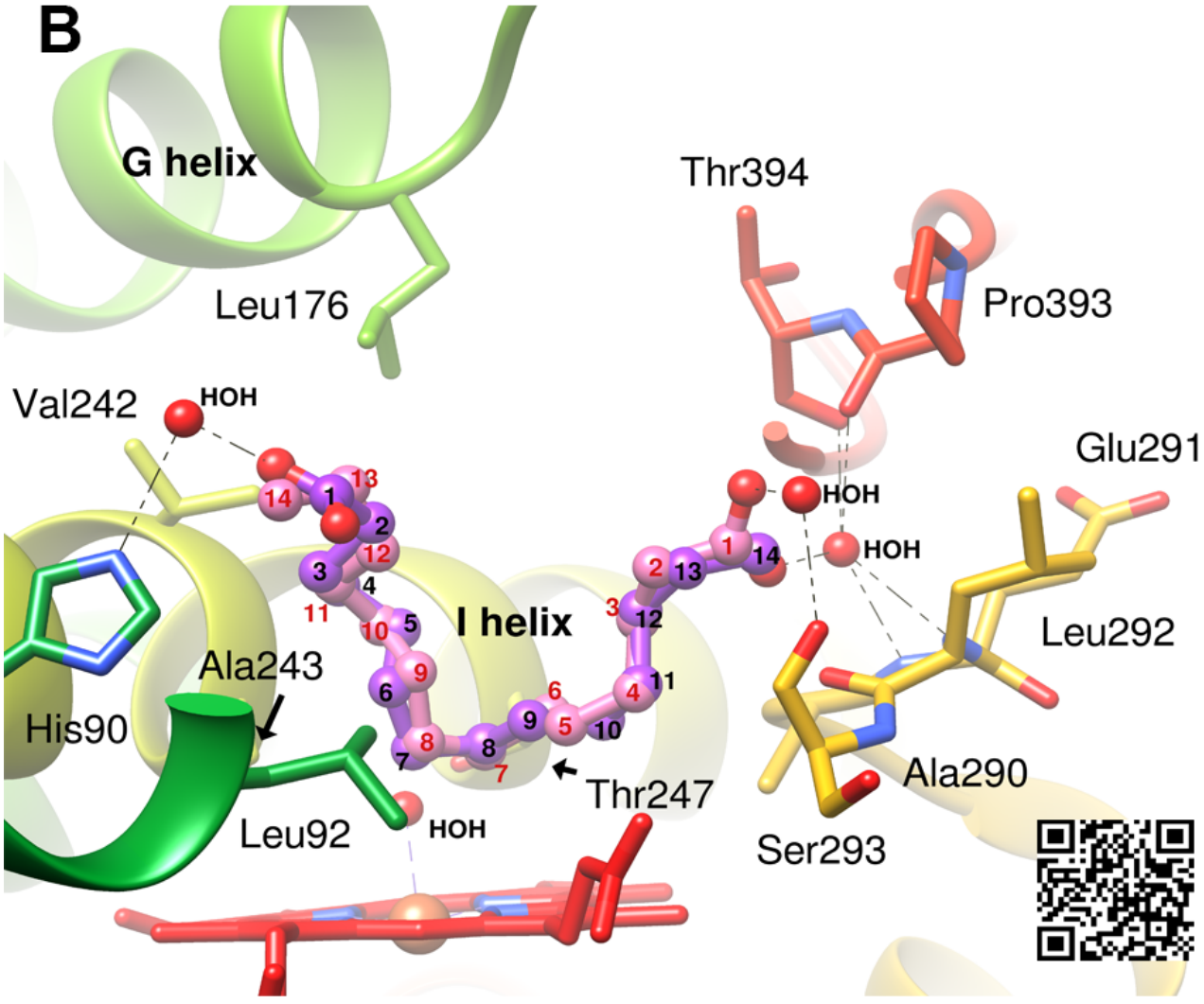

CYP267B1
rel
OH

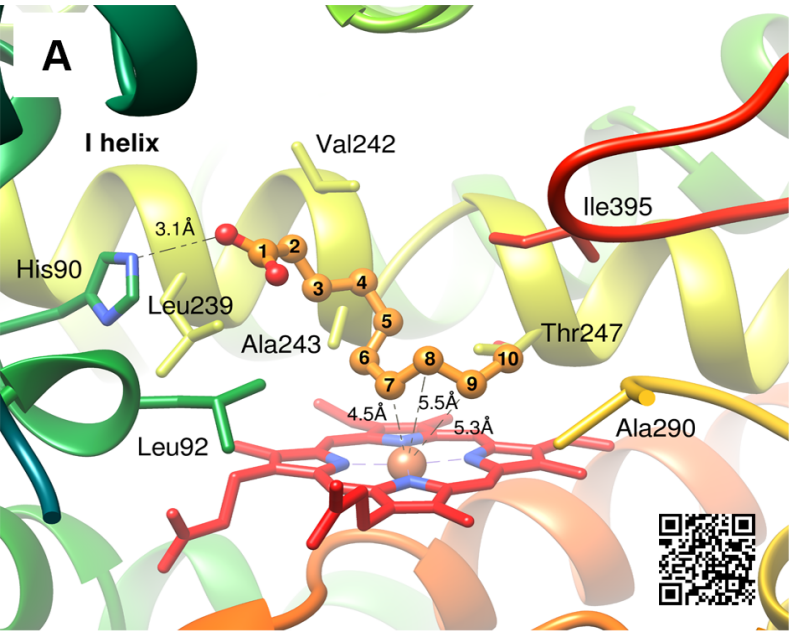

Sheet 3
D
G
F
His6-tag
H
E
L
I
B''
K
C
J
A'
B-C loop
K'
B'
A
B
Sheet 2
Sheet 1

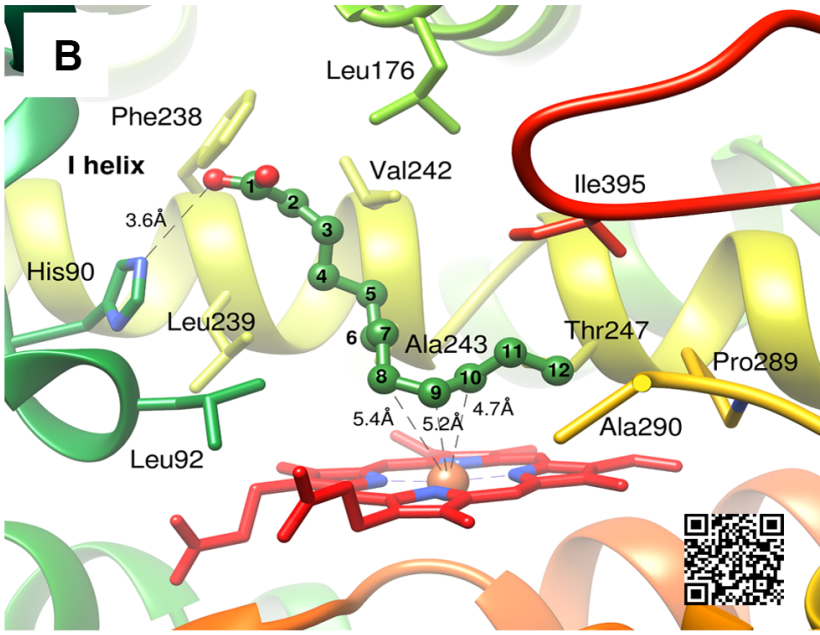

A
G helix
F-G loop
F helix
I helix
B-C loop
H helix

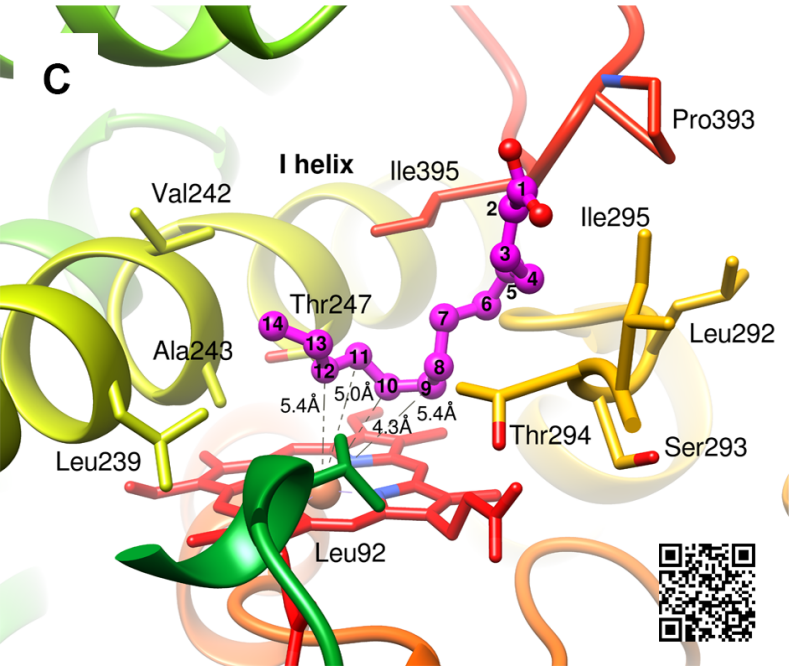

B
Leu80
G helix
F helix
Thr394
Asp88
20Å
16Å
EPOTHILONE D
His90
I helix
9Å
Ala290
Leu92
Cys354

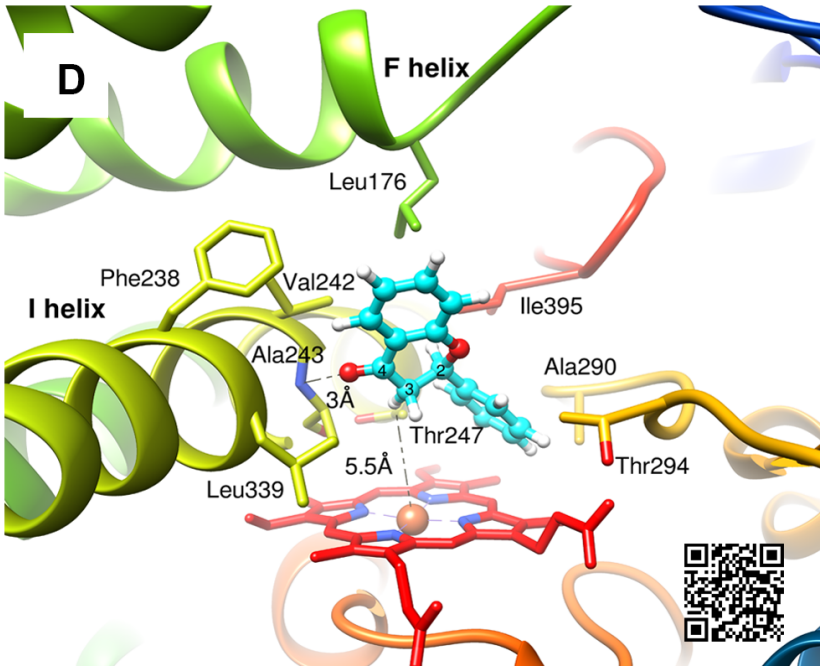

A
I helix
MYR
HOH
HEME

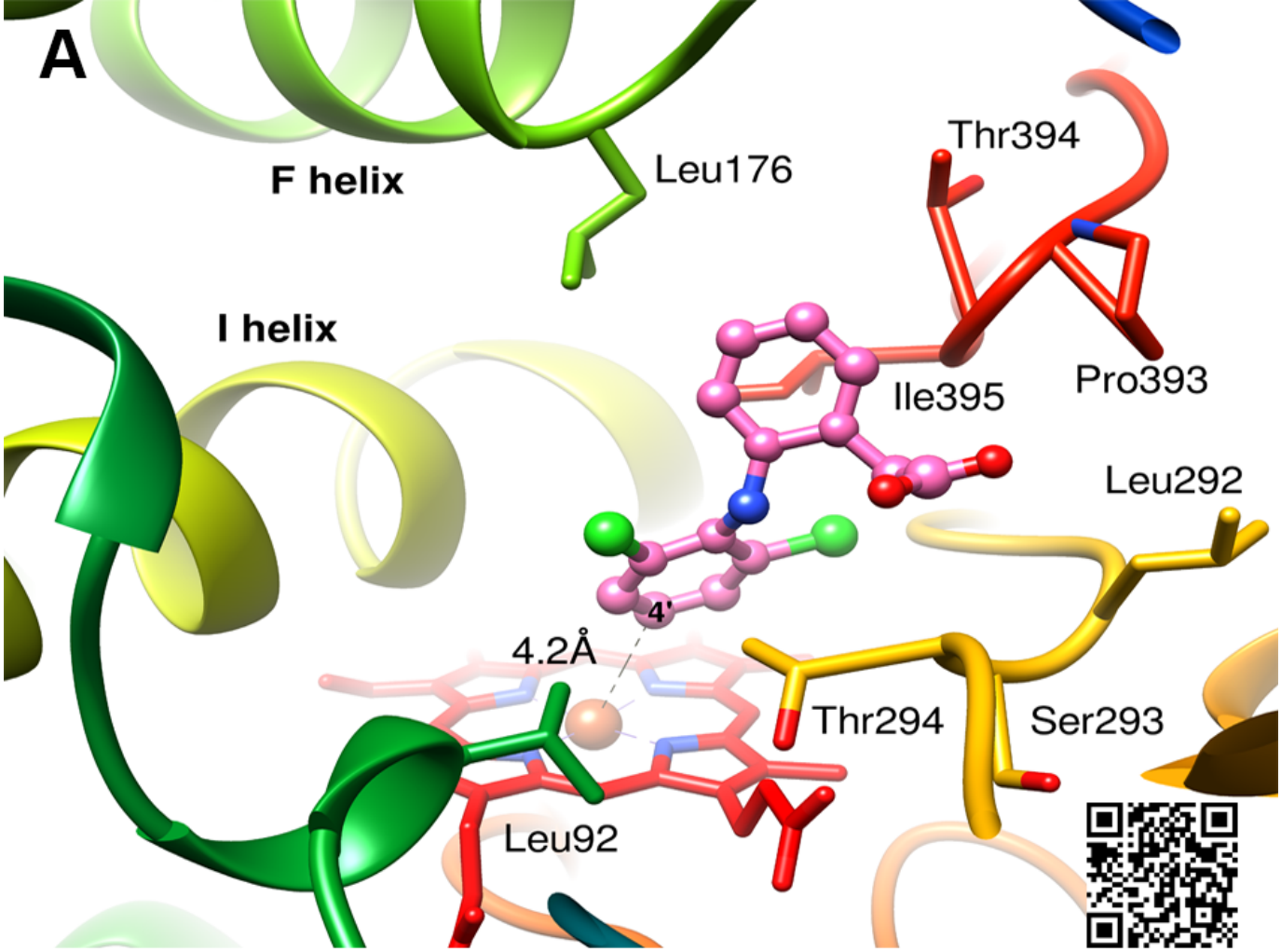

B
G helix
Leu176
Thr394
Pro393
Val242
HOH
Glu291
HOH
HOH
I helix
Leu292
Ala243
His90
Ala290
Thr247
Leu92
HOH
Ser293

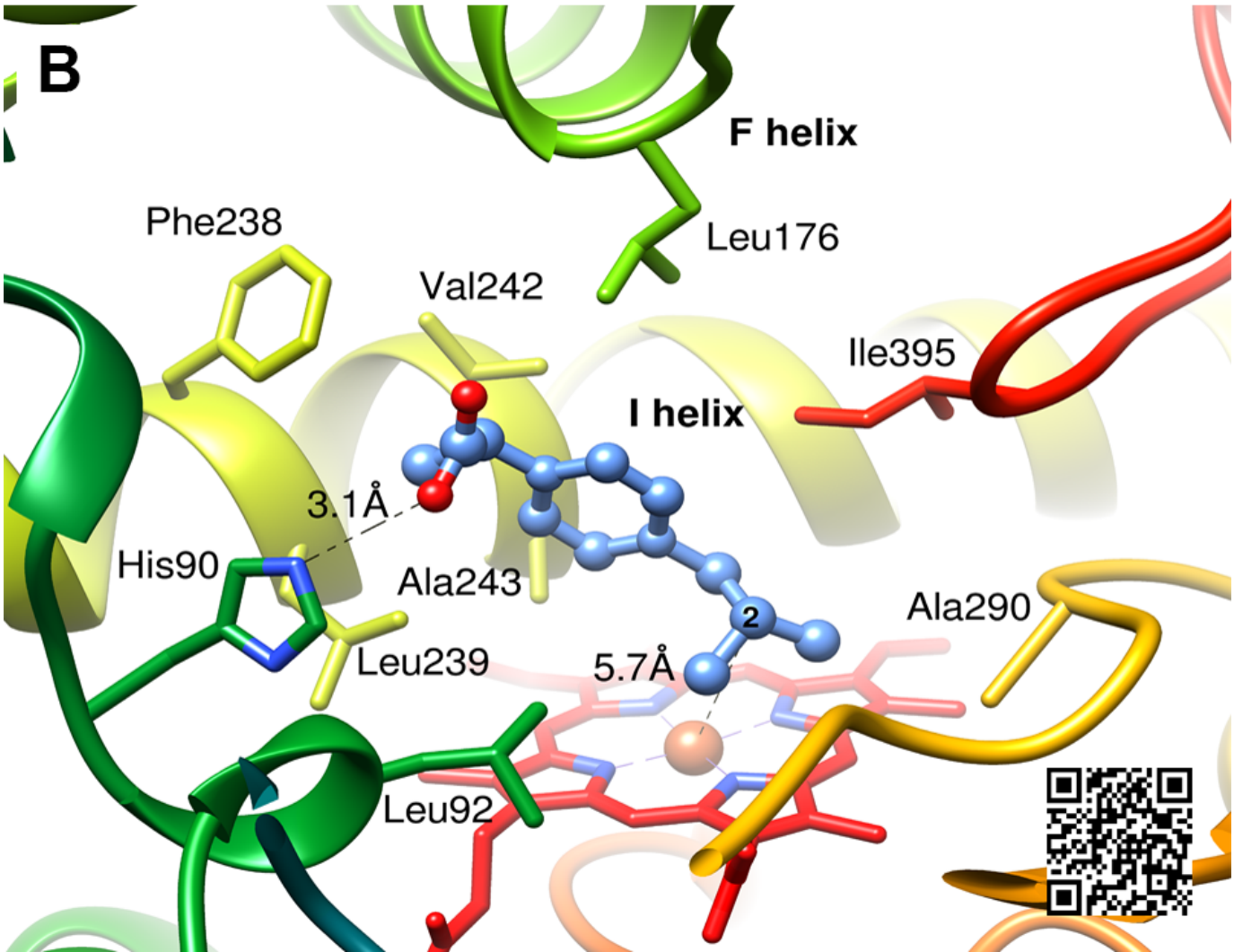

A
I helix
Val242
Ile395
3.1Å
His90
Leu239
Ala243
Thr247
4.5Å
5.5Å
5.3Å
Leu92
Ala290


B
Leu176
Phe238
I helix
Val242
Ile395
3.6Å
His90
Leu239
Ala243
Thr247
Pro289
5.4Å
5.2Å
4.7Å
Ala290
Leu92


C
Pro393
I helix
Ile395
Val242
Ile295
Thr247
Leu292
Ala243
5.4Å
5.0Å
4.3Å
5.4Å
Thr294
Ser293
Leu239
Leu92


D
F helix
Leu176
Phe238
Val242
I helix
Ile395
Ala243
3Å
Ala290
Thr247
5.5Å
Thr294
Leu339


A
F helix
Leu176
Thr394
I helix
Ile395
Pro393
Leu292
4'
4.2Å
Thr294
Ser293
Leu92


B
F helix
Phe238
Leu176
Val242
Ile395
I helix
3.1Å
His90
Ala243
2
Ala290
Leu239
5.7Å
Leu92

Supplementary Information

Structural insights into oxidation of medium-chain fatty acids and flavanone by myxobacterial cytochrome P450 CYP267B1

Ilona K. Jóźwik[1], Martin Litzenburger[2], Yogan Khatri[2Φ], Alexander Schifrin[2], Marco Girhard[3], Vlada Urlacher[3], Andy-Mark W. H. Thunnissen[1 Ψ *] and Rita Bernhardt[2*]

[1] Laboratory of Biophysical Chemistry, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 7, 9747 AG, Groningen, The Netherlands

[2] Department of Biochemistry, Campus B2.2, Saarland University, 66123, Saarbrücken, Germany

[3] Institute of Biochemistry, Heinrich Heine University Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany

[Ψ] Current address: Molecular Enzymology Group, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 4, 9747 AG, Groningen, The Netherlands

[Φ] Current address: University of Michigan, Life Sciences Institute, 210 Washtenaw Ave., Ann Arbor, Michigan 48109, United States

* To whom correspondence should be addressed.

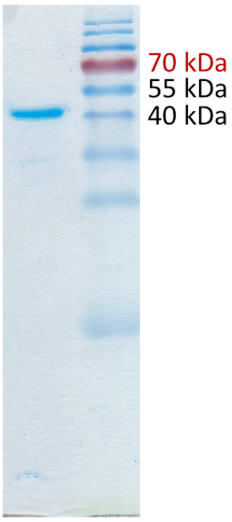
R. Bernhardt, Institute of Biochemistry, Saarland University, Campus B 2.2, 66123 Saarbrücken, Germany. Tel: +49 681 302 4241, Fax: +49 681 302 4739, E-mail: ritabern@mx.uni-saarland.de
or
A.M.W.H. Thunnissen, Molecular Enzymology Group, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 4, 9747 AG, Groningen, The Netherlands. Tel: +31 50 3634380, E-mail: a.m.w.h.thunnissen@rug.nl

Supplementary Table 1: NMR data for the main product of CYP267B1-catalyzed flavanone conversion, the 3-hydroxyflavanone.

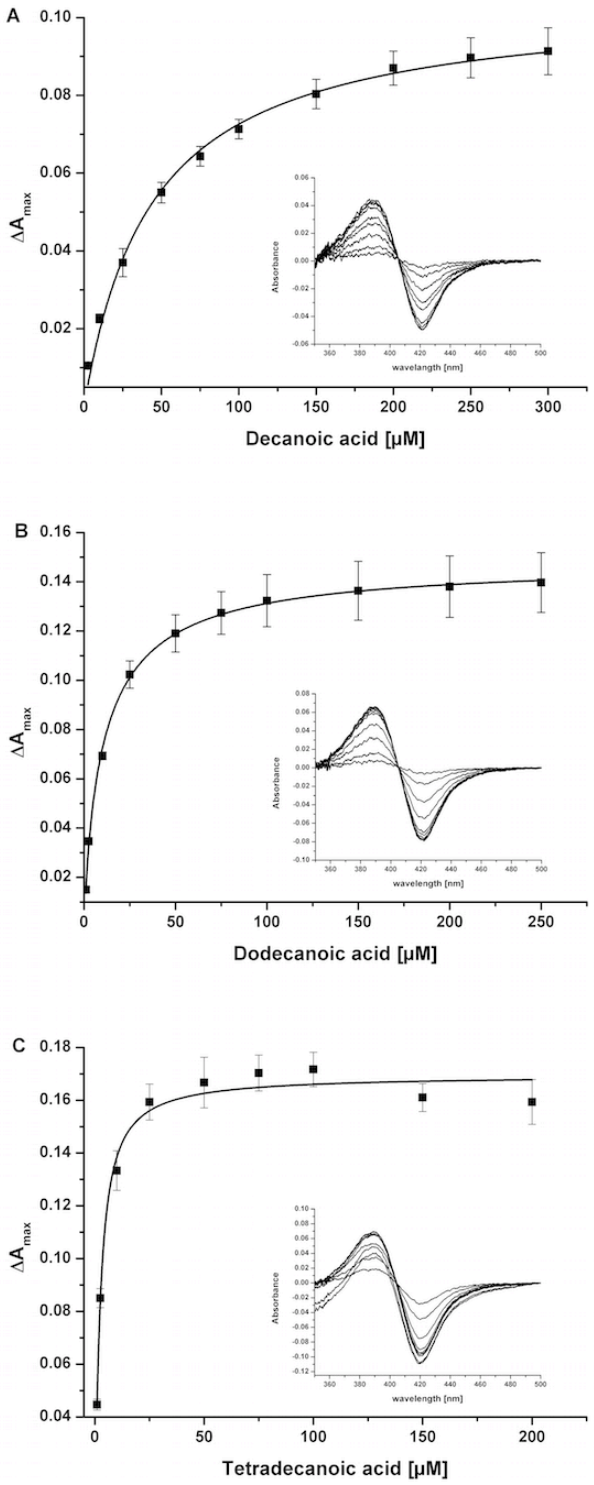
3-hydroxyflavanone (product P1)
^{1}H NMR (500 MHz, $CDCl_3$): 7.91 (m, 1H, H5), 7.58 (m, 2H, H2', H6'), 7.54 (m, 1H, H7), 7.46 (m, 2H, H3', H5'), 7.43 (m, 1H, H4'), 7.10 (m, 1H, H6), 7.04 (d, 1H, H8), 5.13 (d, 1H, H2, J= 12.3 Hz), 4.63 (d, 1H, H3, J= 12.3 Hz)
^{13}C NMR (125 MHz, $CDCl_3$): 194.2 (C4), 161.7 (C8a), 136.9 (C7), 136.3 (C1'), 129.3 (C4'), 128.7 (C3', C5'), 127.5 (C2', C6'),127.3 (C5), 122.1 (C6), 118.1 (C8), 83.9 (C2), 73.6 (C3)

Supplementary Table 2: A summary of the top ten highest-ranked P450 structural homologues of CYP267B1 as evaluated by PDBeFold search [1]. Additionally the result of a search against the PDB entry 2BVJ only is shown in grey. N_{align} – number of matched residues, % Seq – sequence identity.

CYP267B1						
	PDB code and chain id	**Q score**	**RMSD Cα [Å]**	**N_{align}**	**% Seq**	**P450 and organism**
1	4umz:B	0.71	1.52	378	44	PikC (CYP107L1) D50N, *Streptomyces venezuelae*
	2bvj:A	**0.70**	**1.61**	**378**	**44**	**PikC, closed conformation**
2	5cje:A	0.69	1.50	368	43	CYP107L2, *Streptomyces avermitilis*
3	5l92:A	0.68	1.50	364	34	CYP109E1, *Bacillus megaterium*
4	3ejd:H	0.66	1.62	361	40	BioI (CYP107H1), *Bacillus subtilis*
5	3vrm:A	0.63	1.61	362	43	Vdh (CYP107BR1) T107A, *Pseudonocardia autotrophica*
6	3a51:A	0.62	1.61	361	42	Vdh-K1 (CYP107BR1), *Pseudonocardia autotrophica*
7	5x7e:A	0.62	1.74	368	34	CYP105A1-R84A, *Streptomyces griseolus*
8	1t2b:B	0.62	1.55	355	28	P450 CIN (CYP176A1), *Citrobacter braakii*
9	3buj:A	0.62	1.82	369	32	CalO2, *Micromonospora echinospora*
10	4ubs:A	0.62	1.80	365	37	CYP105D7, *Streptomyces avermitilis*
	2bvj:B	**0.57**	**1.71**	**345**	**46**	**PikC, open conformation**

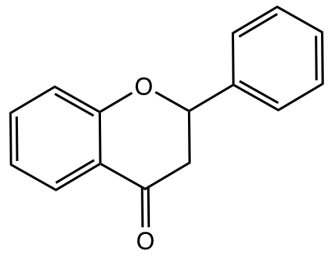


Supplementary Figure 1: Coomassie stained 15 % SDS-PAGE gel of purified CYP267B1.



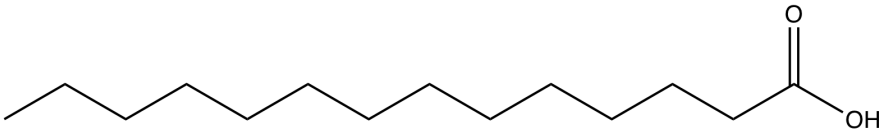
Supplementary Figure 2: Type I spectral shifts induced by the binding of (A) decanoic acid, (B) dodecanoic acid and (C) tetradecanoic acid to CYP267B1. The peak-to-through absorbance differences are plotted against increasing substrate concentrations. The insets show type I spectral shifts induced by the binding of increasing amounts of substrate.

A



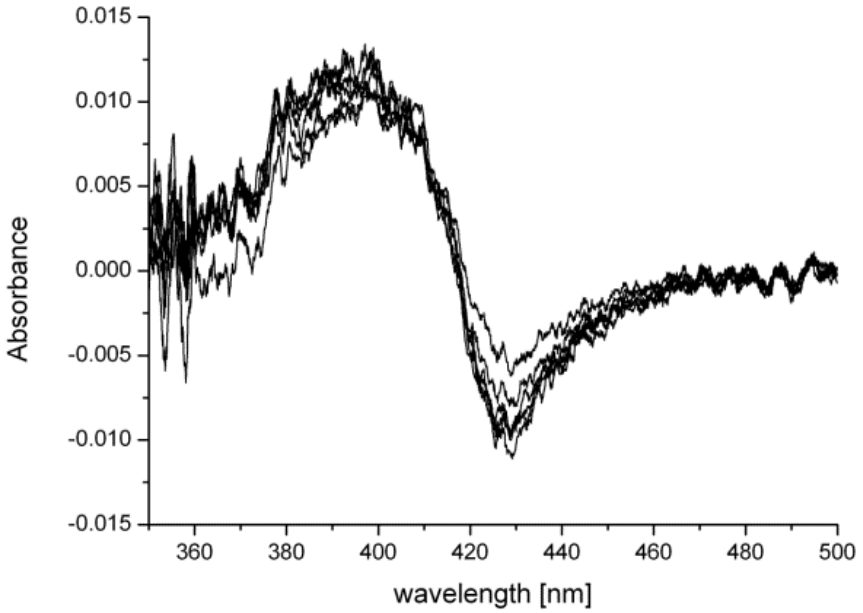
MW = 224.26 g/mol

B



MW = 228.38 g/mol

Supplementary Figure 3: The chemical strutures of flavanone (A) and tetradecanoic acid (B), with their respective molecular masses (MW) given.



Supplementary Figure 4: Spectral perturbations observed upon titration of CYP267B1 with flavanone. Only a minor indication for a characteristic type I spectral change can be observed.

References

1 Krissinel, E. and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr. D Biol. Crystallogr. **60**, 2256-2268